



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 5/10, 1/21, C07K 14/705, 16/28, G01N 33/50, 33/566, C12Q 1/68	A2	(11) International Publication Number: WO 99/67387 (43) International Publication Date: 29 December 1999 (29.12.99)
(21) International Application Number: PCT/US99/14307 (22) International Filing Date: 24 June 1999 (24.06.99) (30) Priority Data: 60/090,579 25 June 1998 (25.06.98) US 09/163,284 29 September 1998 (29.09.98) US 09/261,759 2 March 1999 (02.03.99) US (71) Applicant: MILLENNIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US). (72) Inventor: KHODADOUST, Mehran; 204 Gerry Road, Chestnut Hill, MA 02167 (US). (74) Agents: MANDRAGOURAS, Amy, E.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US) et al.		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: NOVEL MP-7 PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR (57) Abstract <p>Novel MP-7 polypeptides, proteins, and nucleic acid molecules are disclosed. In addition to isolated, full-length MP-7 proteins, the invention further provides isolated MP-7 fusion proteins, antigenic peptides and anti-MP-7 antibodies. The invention also provides MP-7 nucleic acid molecules, recombinant expression vectors containing a nucleic acid molecule of the invention, host cells into which the expression vectors have been introduced and non-human transgenic animals in which a MP-7 gene has been introduced or disrupted. Diagnostic, screening and therapeutic methods utilizing compositions of the invention are also provided.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

NOVEL MP-7 PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

Background of the Invention

5 Signaling factors play an important role in the development and functioning of different cell types by allowing for communication between interacting cells. Such factors provide a signal between cells which can cause cells which recognize the signal to perform specialized tasks, such as cell growth, differentiation and/or proliferation.

For example, cells of the immune system characteristically express a variety of
10 signaling proteins which are crucial to proper functioning of the immune system. Such proteins include secreted immunoglobulins and non-immunoglobulin molecules which interact with cellular adhesion molecules, as well as other selected target molecules. Many of these proteins are members of the immunoglobulin (Ig) superfamily of proteins, characterized by the existence of at least one immunoglobulin (Ig)-like domain.
15 Such proteins function in a variety of immune cell functions ranging from immune cell development and differentiation, antigen recognition, antibody production, cellular signal transduction, and cellular homing of immune responsive cells from the circulation to sites of increased antigen concentration.

Immune cells employ cell surface antigens to facilitate cell-cell interactions
20 which mediate processes such as T-cell activation. A subclass of these cell surface antigens, LY-6 and LY-9, are GPI anchored and mediate various signals between immune cells. The LY-6 subfamily represents multiple cell surface antigenic specificities with distinct cellular and tissue distributions (Dumont *et al.*, 1986, J Immunol 137(1):210-10). LY-6 alloantigens represent a family of phosphatidylinositol
25 anchored proteins that function in the process of T lymphocyte activation. LY-6 is expressed on B lymphocytes, several B cell tumors and bone marrow cells in response to Interferon- γ (IFN- γ) and Tumor Necrosis Factor (TNF). LY-6 is also expressed in thymocytes and T lymphocytes (Malek *et al.*, 1989, 142(6):1929-36). LY-6 can transduce activation signals in T cells and thus may play an important role in T cell
30 function (Dumont *et al.*, 1987, J Immunol 137(8):1183-91). Further, LY-6 mediated T cell activation may be linked to the TCR (T cell receptor) signaling pathway (Codias *et*

- 2 -

al., 1990, J Immunol 145(5):1407-14; (Nickas *et al.*, 1992, Mol Cell Biol 12(1):379-85). LY-6 binding on T lymphocytes leads to stimulation of interleukin-2 production (Malek *et al.*, 1994, Semin Immunol 6(2):105-13), and ligand binding to LY-6C leads to activation of integrins which may facilitate the homing of LY-6C+ CD8+ T cells *in vivo*
5 (Hanninen *et al.*, 1997, PNAS USA 94(13):6898-903).

LY-9 is a lymphocyte cell surface alloantigen expressed in all thymocytes, peripheral lymphocytes and in at least two different T cell functional subsets and B cells (Mathieson *et al.*, 1980, J Immunol 125(5):2127-36). LY-9 belongs to the subgroup of the Ig superfamily that includes Bcm-1, CD2, and LFA-3. (Sandrin *et al.*, 1992, J
10 Immunol. 149(5):1636-41). LY-9 may participate in adhesion reactions between T lymphocytes and accessory cells by homophilic interaction. (Kingsmore *et al.*, 1995, Immunogenetics 42(1):59-62). Additionally, LY-9 is expressed on the cell surface of many thymomas and, thus, may serve as a marker for identifying lymphocyte-derived tumors. (Hogarth *et al.*, 1982, J Natl Cancer Inst 69(3):619-26).

15 Given the importance of such signaling proteins in the proper functioning of the body, there exists a need to identify novel signaling factors which function to regulate the systems such as the immune response and whose aberrant function can result in disorders arising from improper cell-cell interactions such as immune disorders and cardiovascular disorders, e.g., congestive heart failure, ischemia, cardiac hypertrophy or
20 a disorder arising from improperly regulated ankyrin repeat containing protein action on target molecules/cells giving rise to improperly regulated cellular processes.

Summary of the Invention

The present invention is based, at least in part, on the discovery of novel cell-
25 surface molecules, referred to herein as "Myocardium Protein -7" ("MP-7") nucleic acid and protein molecules. The MP-7 molecules of the present invention are useful as modulating agents in regulating a variety of cellular processes. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding MP-7 proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as
30 primers or hybridization probes for the detection of MP-7-encoding nucleic acids.

- 3 -

In one embodiment, a MP-7 nucleic acid molecule includes a nucleotide sequence at least about 25%, 30%, 35%, 40%, 43%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a complement thereof. In a preferred embodiment, the isolated nucleic acid molecule includes a nucleotide sequence shown in SEQ ID NO:3 or a complement thereof. In another embodiment, the nucleic acid molecule further comprises nucleotides 1006-1008 of SEQ ID NO:1. In another preferred embodiment, an isolated nucleic acid molecule has the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3 or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a complement thereof. In yet another preferred embodiment, an isolated nucleic acid molecule has the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a complement thereof.

In another embodiment, a MP-7 nucleic acid molecule includes a nucleotide sequence encoding a protein or polypeptide having an amino acid sequence sufficiently homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4. In a preferred embodiment, a MP-7 nucleic acid molecule includes a nucleotide sequence encoding a protein or polypeptide which includes an amino acid sequence at least 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4. In yet another embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein or polypeptide having an amino acid sequence of at least 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homology to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4. In another preferred embodiment, an isolated nucleic acid molecule encodes the amino acid sequence of human MP-7. In yet another preferred embodiment, the nucleic acid molecule includes a nucleotide sequence encoding a protein having the amino acid sequence of SEQ ID NO: 2, SEQ ID NO:4, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number 209887.

- 4 -

In another embodiment, an isolated nucleic acid molecule of the present invention encodes a protein, preferably a MP-7 protein, which includes at least one and, preferably, two or more Ig-like domains. In another embodiment, an isolated nucleic acid molecule of the present invention encodes a protein, preferably a MP-7 protein, which includes a first and second extracellular Ig-like domain. In another embodiment, an isolated nucleic acid molecule of the present invention encodes a protein, preferably a MP-7 protein, which includes a transmembrane domain. In another embodiment, an isolated nucleic acid molecule of the present invention encodes a protein, preferably a MP-7 protein, which includes a cytoplasmic domain. In another embodiment, an isolated nucleic acid molecule of the present invention encodes a protein, preferably a MP-7 protein, which includes a first and second Ig-like domain and a transmembrane domain. In another embodiment, an isolated nucleic acid molecule of the present invention encodes a protein, preferably a MP-7 protein, which includes a first and second Ig-like domain, a transmembrane domain and a cytoplasmic domain and, preferably, is localized to the cell surface. In yet another embodiment, a MP-7 nucleic acid molecule encodes a MP-7 protein and is a naturally occurring nucleotide sequence.

Another embodiment of the invention features nucleic acid molecules, preferably MP-7 nucleic acid molecules, which specifically detect MP-7 nucleic acid molecules relative to nucleic acid molecules encoding non-MP-7 proteins. For example, in one embodiment, such a nucleic acid molecule is at least 100, preferably 100-200, more preferably 200-300, more preferably 300-400, and even more preferably 400-417 nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3 the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a complement thereof. In preferred embodiments, the nucleic acid molecules are at least 15, 18, 20, 22, 35, 40, 50 (e.g., contiguous) nucleotides in length and hybridize under stringent conditions to nucleotides 418-499 of SEQ ID NO:1, SEQ ID NO:3. In other preferred embodiments, the nucleic acid molecules comprise nucleotides 418-499 of SEQ ID NO:1, SEQ ID NO:3.

- 5 -

In other preferred embodiments, the nucleic acid molecule encodes a naturally occurring allelic variant of a polypeptide which includes the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule which includes SEQ ID NO:1 or SEQ ID NO:3 under stringent conditions.

Another embodiment of the invention provides an isolated nucleic acid molecule which is antisense to a MP-7 nucleic acid molecule, e.g., the coding strand of a MP-7 nucleic acid molecule.

Another aspect of the invention provides a vector comprising a MP-7 nucleic acid molecule. In certain embodiments, the vector is a recombinant expression vector. In another embodiment, the invention provides a host cell containing a vector of the invention. The invention also provides a method for producing a protein, preferably a MP-7 protein, by culturing in a suitable medium, a host cell, e.g., a mammalian host cell such as a non-human mammalian cell, of the invention containing a recombinant expression vector such that the protein is produced.

Another aspect of this invention features isolated or recombinant MP-7 proteins and polypeptides. In one embodiment, an isolated protein, preferably a MP-7 protein, includes at least one and, preferably, two or more Ig-like domains. In another embodiment, an isolated protein, preferably a MP-7 protein, includes a first and second extracellular Ig-like domain. In another embodiment, an isolated protein, preferably a MP-7 protein, includes a transmembrane domain. In another embodiment, an isolated protein, preferably a MP-7 protein, includes a cytoplasmic domain. In another embodiment, an isolated protein, preferably a MP-7 protein, includes a first and second Ig-like domain and a transmembrane domain. In another embodiment, an isolated protein, preferably a MP-7 protein, includes a first and second Ig-like domain, a transmembrane domain, and a cytoplasmic domain and, preferably, is localized to the cell surface. In another embodiment, an isolated protein, preferably a MP-7 protein, has an amino acid sequence sufficiently homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4.

- 6 -

In a preferred embodiment, a protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4. In another preferred

5 embodiment, a protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 and has an Ig-like domain.

In another preferred embodiment, a protein or polypeptide, preferably a MP-7
10 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 and has a first and second Ig-like domain. In another preferred embodiment, a protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%,
15 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 and has a transmembrane domain.

In yet another preferred embodiment, a protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%,
20 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 and has a cytoplasmic domain. In a preferred embodiment, a protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more
25 homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, and has an Ig-like domain and a transmembrane domain.

In a preferred embodiment, a protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the
30 amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, and has a first and second extracellular Ig-like domain and a transmembrane domain. In a preferred embodiment, a

- 7 -

protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, and has a first and second Ig-like domain and a cytoplasmic domain. In yet another embodiment, a protein or polypeptide, preferably a MP-7 protein, includes an amino acid sequence at least about 25%, 30%, 31%, 32%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, and has a transmembrane domain and a cytoplasmic domain.

10 In another embodiment, the invention features fragments of the proteins having the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, wherein the fragment comprises at least about 10, 12, 15, 20, 30 or more amino acids (e.g., contiguous amino acids) of the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, or an amino acid or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession No. 209887. In another embodiment, a protein, preferably a MP-7 protein, includes the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4. In yet another embodiment, the protein has the amino acid sequence SEQ ID NO:2, SEQ ID NO:4.

Another embodiment of the invention features an isolated protein, preferably a MP-7 protein, which is encoded by a nucleic acid molecule which includes a nucleotide sequence at least about 25%, 30%, 35%, 40%, 42%, 43%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to a nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or a complement thereof. This invention further features an isolated protein, preferably a MP-7 protein, which is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or a complement thereof.

The proteins of the present invention, preferably MP-7 proteins, or biologically active portions thereof, can be operatively linked to a non-MP-7 polypeptide (e.g., heterologous amino acid sequences) to form fusion proteins. The invention further features antibodies, such as monoclonal or polyclonal antibodies, that specifically bind

- 8 -

proteins of the invention, preferably MP-7 proteins. In addition, the MP-7 proteins or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides a method for detecting MP-7
5 expression in a biological sample by contacting the biological sample with an agent capable of detecting a MP-7 nucleic acid molecule, protein or polypeptide such that the presence of a MP-7 nucleic acid molecule, protein or polypeptide is detected in the biological sample.

In another aspect, the present invention provides a method for detecting the
10 presence of MP-7 activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of MP-7 activity such that the presence of MP-7 activity is detected in the biological sample.

In another aspect, the invention provides a method for modulating MP-7 activity comprising contacting a cell capable of expressing MP-7 with an agent that modulates
15 MP-7 activity such that MP-7 activity in the cell is modulated. In one embodiment, the agent inhibits MP-7 activity. In another embodiment, the agent stimulates MP-7 activity. In one embodiment, the agent is an antibody that specifically binds to a MP-7 protein. In another embodiment, the agent modulates expression of MP-7 by modulating transcription of a MP-7 gene or translation of a MP-7 mRNA. In yet another
20 embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of a MP-7 mRNA or a MP-7 gene.

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant MP-7 protein or nucleic acid expression or activity by administering an agent which is a MP-7 modulator to the
25 subject. In one embodiment, the MP-7 modulator is a MP-7 protein. In another embodiment the MP-7 modulator is a MP-7 nucleic acid molecule. In yet another embodiment, the MP-7 modulator is a peptide, peptidomimetic, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant MP-7 protein or nucleic acid expression is a immune disorder, a cardiovascular disorder, e.g., congestive
30 heart failure, cardiomyopathy, or a disorder arising from improper cell-cell or cell-ligand

In one embodiment, the methods of the present invention are used to treat a subject having a disorder characterized by aberrant MP-7 protein or nucleic acid expression or activity by administering an agent which is a MP-7 modulator to the subject. In one embodiment, the MP-7 modulator is a MP-7 protein. In another
5 embodiment the MP-7 modulator is a MP-7 nucleic acid molecule. In yet another embodiment, the MP-7 modulator is a peptide, peptidomimetic, or other small molecule. In a preferred embodiment, the disorder characterized by aberrant MP-7 protein or nucleic acid expression is a cell-cell interaction based disorder.

The present invention also provides a diagnostic assay for identifying the
10 presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a MP-7 protein; (ii) mis-regulation of said gene; and (iii) aberrant post-translational modification of a MP-7 protein, wherein a wild-type form of said gene encodes an protein with a MP-7 activity.

In another aspect the invention provides a method for identifying a compound
15 that binds to or modulates the activity of a MP-7 protein, by providing a indicator composition comprising a MP-7 protein having MP-7 activity, contacting the indicator composition with a test compound, and determining the effect of the test compound on MP-7 activity in the indicator composition to identify a compound that modulates the activity of a MP-7 protein.

20 Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1 depicts the cDNA sequence of human MP-7. The nucleotide sequence
25 corresponds to nucleic acids 1 to 1008 of SEQ ID NO:1, SEQ ID NO:3.

Figure 2 depicts a predicted amino acid sequence of human MP-7. The amino acid sequence correspond to amino acids 1 to 335 of SEQ ID NO:2.

Figure 3 depicts the alignment between the MP-7 protein and the human LY-9 protein. This alignment were generated utilizing the DNASTAR® Megalign program
30 Version 3.14 and specifically the Clustal utility with the following parameter settings: a PAM250 weight residue table, a gap length penalty of 10, and a gap penalty of 10.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as MP-7 nucleic acid and polypeptide molecules, which
5 play a role in or function in a variety of cellular processes, e.g., cardiac cellular processes. In one embodiment, the MP-7 molecules modulate the activity of one or more proteins involved in a cardiovascular disorder, e.g., congestive heart failure, cardiomyopathy. In another embodiment, the MP-7 molecules of the present invention are capable of modulating the transcription of genes involved in a cardiovascular
10 disorder, e.g., congestive heart failure, cardiomyopathy. In one embodiment, the MP-7 molecules modulate the activity of one or more proteins involved in an immune cell disorder, e.g., thymoma. In another embodiment, the MP-7 molecules of the present invention are capable of modulating the transcription of genes involved in an immune cell disorder, e.g., thymoma.

15 As used herein, the term "cardiovascular disorder" includes a disease, disorder, or state involving the cardiovascular system, e.g., the heart, the blood vessels, and/or the blood. A cardiovascular disorder can be caused by an imbalance in arterial pressure, a malfunction of the heart, or an occlusion of a blood vessel, e.g., by a thrombus. Examples of such disorders include hypertension, atherosclerosis, coronary artery
20 spasm, coronary artery disease, valvular disease, arrhythmias, and cardiomyopathies.

As used herein, the term "congestive heart failure" includes a condition characterized by a diminished capacity of the heart to supply the oxygen demands of the body. Symptoms and signs of congestive heart failure include diminished blood flow to the various tissues of the body, accumulation of excess blood in the various organs, e.g.,
25 when the heart is unable to pump out the blood returned to it by the great veins, exertional dyspnea, fatigue, and/or peripheral edema, e.g., peripheral edema resulting from left ventricular dysfunction. Congestive heart failure may be acute or chronic. The manifestation of congestive heart failure usually occurs secondary to a variety of cardiac or systemic disorders that share a temporal or permanent loss of cardiac function.
30 Examples of such disorders include hypertension, coronary artery disease, valvular disease, and cardiomyopathies, e.g., hypertrophic, dilative, or restrictive

- 11 -

cardiomyopathies. Congestive heart failure is described in, for example, Cohn J.N. *et al.* (1998) *American Family Physician* 57:1901-04, the contents of which are incorporated herein by reference.

As used herein, the term "cardiomyopathy" is art-recognized and includes

5 conditions characterized by myocardial disorders with impaired ventricular function. Symptoms and signs of cardiomyopathy include congestive cardiomyopathy, conduction defects, atrial fibrillation or flutter, ventricular arrhythmia, congestive heart failure, and pericardial effusion. Typically arrhythmias are the earliest manifestation of cardiomyopathy, becoming sustained by the third or fourth decade of life with

10 progressive cardiomyopathy becoming clinically apparent in the fourth and fifth decade. In the late stage, the fifth and sixth decade of life, cardiomyopathy typically manifests as chamber dilation and systolic dysfunction. Sudden death can also occur in the late stage. Examples of such disorders include hypertrophic, dilative, and restrictive cardiomyopathies. Cardiomyopathy is described in, for example,

15 <http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispim?115200.cs>, and Kass, S. *et al.*, "A gene defect that causes conduction system disease and dilated cardiomyopathy maps chromosome 1p1-1q1," *Nature Genet.* 7:546-551. 1994, the contents of which are incorporated herein by reference.

As used herein, the term "cardiac cellular processes" includes intra-cellular or

20 inter-cellular processes involved in the functioning of the heart. Cellular processes involved in the nutrition and maintenance of the heart, the development of the heart, or the ability of the heart to pump blood to the rest of the body are intended to be covered by this term. Such processes include, for example, cardiac muscle contraction, distribution and transmission of electrical impulses, and cellular processes involved in

25 the opening and closing of the cardiac valves. The term "cardiac cellular processes" further includes processes such as the transcription, translation and post-translational modification of proteins involved in the functioning of the heart, e.g., myofilament specific proteins, such as troponin I, troponin T, myosin light chain 1 (MLC1), and α -actinin.

- 12 -

The present invention is based on the discovery of novel molecules, referred to herein as MP-7 protein and nucleic acid molecules, which comprise a family of molecules having certain conserved structural and functional features. The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin. Members of a family may also have common functional characteristics.

The MP-7 nucleic acid molecules encode polypeptides, referred to herein as MP-7 polypeptides. In one embodiment, MP-7 polypeptides of the invention are involved in cell surface signaling. In a preferred embodiment, the MP-7 polypeptides of the invention are involved in the regulation of transcription factors which are involved in cell surface signaling.

The MP-7 nucleic acid molecule and polypeptides have sequence similarity with the transmembrane glycosyl phosphatidylinositol (GPI) anchored LY-9 protein which is found on thymocytes, bone marrow progenitor cells and mature T and B cells, but not blast cells which have committed to a red blood cell phenotype or granulocyte maturation pathway (Miller *et al.*, 1985, *J. Immunol.* 134:3286). Accordingly, MP-7 polypeptides of the invention may interact with (e.g., bind to) at least one ligand which is a extracellular signalling molecule and, thus, may be involved in the regulation of immune cell, e.g., T and B cell, activation and differentiation.

The MP-7 gene has been mapped to the region near the CMD1A gene, cardiomyopathy, dilated, 1A. Disorders arising from mutations in this gene have been typified by a heterogenous group of primary myocardial disorders with impaired ventricular function (<http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispim?115200>). Accordingly, MP-7 polypeptides of the invention may interact with (e.g., bind to) at least one ligand which is involved in cardiomyopathy, e.g., the maintenance of atrioventricular conduction, and the maintenance of cardiac contractility.

- 13 -

MP-7 family members are identified based on the presence of at least one and, preferably, two, three, four or more Ig-like domains in the protein or corresponding nucleic acid molecule. As used herein, the term "immunoglobulin (Ig)-like domain" includes a protein domain having an amino acid sequence of at least about 40-90, preferably at least about 50-75 or 80, more preferably at least about 53 or 73 amino acid residues in length of which at least about 30%, preferably at least about 40%, more preferably at least about 50, 60 or 70% of the amino acids are homologous to the amino acid sequence of an immunoglobulin domain e.g., the immunoglobulin domain of antibodies (Huber, 1980, *Klin Wochenscher* 58(22):1217-31; Mellman et al., 1988, *J Cell Sci Suppl* 9:45-65), the giant muscle kinase Titin (Labeit et al., 1997, *Circ Res* 80(2):290-294) and the receptor tyrosine kinases (Yao et al., 1995, *J Immunol* 155(2):652-661; Tseng et al., *Immunol Res* 13(4):299-310).

Immunoglobulin-like domains may be involved in protein-protein and protein-ligand interactions. The homologous amino acids between an Ig-like domain and an Ig domain can be positioned across the entire Ig-like domain, or, alternatively, the homologous amino acids between an Ig-like domain and an Ig domain can be concentrated in regions of high homology dispersed throughout the Ig-like domain. Additionally, the term an "immunoglobulin-like domain" includes amino acid sequences having a bit score for the alignment of the sequence to the Ig family Hidden Markov Model (HMM) of at least 5, preferably 5-10, more preferably 10-15, more preferably 15-20, even more preferably 20-25, 25-55, 55-100 or greater. The Ig family HMM has been assigned the PFAM Accession PF00047 (<http://genome.wustl.edu/Pfam/WWWdata/ig.html>).

To identify the presence of an Ig-like domain in a MP-7 family member, the amino acid sequence of the family member is searched against a database of HMMs (e.g., the Pfam database, release 2.1) using the default parameters (http://www.sanger.ac.uk/Software/Pfam/HMM_search). For example, the hmmsf program, which is available as part of the HMMER package of search programs, is a family specific default program for PF00047 having a score of 5 as the default threshold score for determining a hit. For example, a search using the amino acid sequence of SEQ ID NO:2 was performed against the HMM database resulting in the identification

- 14 -

of an Ig-like domain in the amino acid sequence of SEQ ID NO:2 and a score of 10.27 against the Ig family HMM Accession PF00047. The results of the search are set forth below.

5 Score: 10.27 SEQ ID NO:2: aa34-107 HMM: aa1-47

```

          *GqsVTLTCmVs fhPpdYt .IwWYrNgqpi.....
34    GGA VTFPLKSK--VKQVDSIVWTFNTTPLVTIQPEGGTIIVTQNRNRER

10                    .....tLtInsWqyEDsGtYwCmV*
81    VDFPDGGYSLKLSKLKKNDSGIYYVGI        107

```

In another example, a search was performed using the amino acid sequence of SEQ ID NO:2 against the HMM database resulting in the identification of an Ig-like domain in the amino acid sequence of SEQ ID NO:2 and a score of 6.60 against the Ig family HMM Accession PF00047. The results of the search are set forth below.

Score: 6.60 SEQ ID NO:2: aa144-197 HMM: aa1-47

```

20            *GqsVTLTCmVs fhPpdYtIwWYrNgqpi.....tLtInsWqyEDs.Gt
144    TCVTNLTCCMEHGEEDEVITYWKALGQAANESHNGSILPISWRWGESDMT

          YwCmV*
25    193 FICVA            197

```

Accordingly, in one embodiment, a MP-7 protein includes a first Ig-like domain of about 70-75, preferably about 73 amino acid residues and, preferably, includes a second Ig-like domain of about 50-55, preferably, about 53 amino acid residues. In a preferred embodiment, a first and second Ig-like domains are located in the N-terminal or extracellular region of a MP-7 protein. For example, in one embodiment, a MP-7 protein contains a first Ig-like domain of about amino acids 34-107 of SEQ ID NO:2,

- 15 -

SEQ ID NO:4, and a second Ig-like domain of about amino acids 144-197 of SEQ ID NO:2, SEQ ID NO:4.

In another embodiment, MP-7 family members include at least one Ig-like domain having at least about 1, 2, 3, 4 or more Protein kinase C (PKC) phosphorylation sites. PKC phosphorylation sites can be found at least within the amino acid sequence 1-226 of SEQ ID NO:2, SEQ ID NO:4. The Ig-like domain can further include at least about one Casein kinase II phosphorylation sites. Casein kinase II phosphorylation sites can be found at least within the amino acid sequence 1-226 of SEQ ID NO:2, SEQ ID NO:4. The Ig-like domain can further include at least about one Tyrosine kinase phosphorylation site. Tyrosine kinase phosphorylation sites can be found at least within the amino acid sequence 1-226 of SEQ ID NO:2, SEQ ID NO:4. The Ig-like domain further may include at least about 2-9, 3-8, 4-7 and preferably 6 N-glycosylation sites. N-glycosylation sites can be found at least within the amino acid sequence 1-226 of SEQ ID NO:2, SEQ ID NO:4. The Ig-like domain can further include at least about 4-10, 5-9, 6-8 and preferably 7 N-myristoylation sites. N-myristoylation sites can be found at least within the amino acid sequence 1-226 of SEQ ID NO:2, SEQ ID NO:4.

In another embodiment, MP-7 family members are identified based on the presence of at least one leucine zipper pattern or domain. As used herein, a "leucine zipper domain" includes, at least, an amino acid motif having an amino acid sequence of about 5-40, preferably about 10-35, and more preferably 15-30, 20-25, or 22 amino acid residues in length. In one embodiment, the leucine zipper domain has a leucine zipper consensus sequence: L-X (6)-L-X (6)-L-X (6)-L, corresponding to SEQ ID NO:4. Accordingly, in one embodiment, a MP-7 protein is human MP-7 having a leucine zipper domain of about 22 amino acid residues. A leucine zipper domain can be found at least at amino acids 229-250 of SEQ ID NO:2, SEQ ID NO:4. The leucine zipper consensus sequence is further described in PROSITE Document, Accession No. PDOC00029 (<http://expasy.hcuge.ch/cgi-bin/get-prodoc-entry?PDOC00029>) and as PROSITE Accession No. PS00029.

The domains described herein are described according to standard Prosite Signature designation (e.g., all amino acids are indicated according to their universal single letter designation; X designates any amino acid; X (n) designates an alphanumeric

- 16 -

number of "n" amino acids, e.g., X (2) designates any 2 amino acids; and [LIVM] indicates any one of the amino acids appearing within the brackets, e.g., any one of L, I, V, or M, in the alternative, any one of Leu, Ile, Val, or Met.)

In another embodiment of the invention, a MP-7 family member is identified based on the presence of a transmembrane domain. As used herein, a "transmembrane domain" includes a protein domain having an amino acid sequence containing at least about 5, preferably about 10, more preferably about 15-30, 20-25 or, preferably, about 23 or 24 amino acid residues, of which at least about 60-70%, preferably about 80% and more preferably about 90% of the amino acid residues contain non-polar side chains, for example, alanine, valine, leucine, isoleucine, proline, phenylalanine, tryptophan, and methionine. A transmembrane domain is lipophilic in nature. In one embodiment, the transmembrane domain has at least about 65-95%, preferably at least about 70-90%, and more preferably at least about 75-85% homology with the amino acid sequence of a transmembrane domain of the human MP-7 amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4. In another embodiment, a transmembrane domain has amino acid residues 226-250 of the amino acid sequence as set forth in SEQ ID NO:2, SEQ ID NO:4. As further defined herein, a transmembrane domain of a MP-7 protein family member, however, is not sufficiently homologous to the amino acid sequence of a member of another protein family, such as a non-MP-7 protein family.

In another embodiment of the invention, a MP-7 family member is identified based on the presence of a cytoplasmic domain. As used herein, a "cytoplasmic domain" refers to a protein domain which is at least about 65-105 amino acid residues in length, preferably at least about 70-100 amino acid residues in length, and more preferably at least about 75-95, or at least about 80-90 or 85 amino acid residues in length, and has at least about 65-95%, preferably at least about 70-90%, and more preferably at least about 75-85% homology with the amino acid sequence of a cytoplasmic domain of a human MP-7 sequence set forth in SEQ ID NO:2, SEQ ID NO:4. In another embodiment, a cytoplasmic domain has amino acid residues 250-335 of the amino acid sequence as set forth in SEQ ID NO:2, SEQ ID NO:4. As further defined herein, a cytoplasmic domain of a MP-7 protein family member, however, is not sufficiently homologous to the amino acid sequence of a member of another protein family, such as a non-MP-7 protein family.

In a preferred embodiment, the cytoplasmic domain of an MP-7 family member contains at least one, and preferably two Protein kinase C (PKC) phosphorylation sites. PKC phosphorylation sites can be found at least within the amino acid sequence 250-335 of SEQ ID NO:2, SEQ ID NO:4. In addition, the cytoplasmic domain preferably further
5 contains at least about 2-4 and preferably 3 Casein kinase II phosphorylation sites. Casein kinase II phosphorylation sites can be found at least within the amino acid sequence 250-335 of SEQ ID NO:2, SEQ ID NO:4. In a preferred embodiment, the cytoplasmic domain further contains at least about one N-glycosylation site. N-glycosylation sites can be found at least within the amino acid sequence 250-335 of SEQ ID NO:2, SEQ ID NO:4.

10 In another embodiment, the MP-7 family members of the invention are identified based in the presence of at least one and, preferably more of the following sites: at least one PKC phosphorylation site, at least one Tyrosine kinase phosphorylation site, at least one N-glycosylation site, at least one N-myristoylation site, and at least one casein kinase II phosphorylation site, wherein the site(s) have a consensus sequence selected from: [ST]-X-
15 [RK], where S or T is a phosphorylation site; [RK]-X (2)-[DE]-X (3)-Y (see PROSITE document for alternative consensus sequences); N-{P}-[ST]-{P}, where N is a glycosylation site; G-{EDRKHPFYW}-X (2)-[STAGCN]-{P}, where G is an N-myristoylation site; and [ST]-X (2)-[DE], where S or T is a phosphorylation site, respectively. X designates any amino acid; (n) designates an alphanumeric number of "n"
20 amino acids. These sites are further described in PROSITE Documents, Accession No. PDOC00005, PDOC00007, PDOC00001, PDOC00008, PDOC00006, respectively (<http://expasy.hcuge.ch/cgi-bin/get-prodoc-entry?PDOC00005>, PDOC00007, PDOC00001, PDOC00008, PDOC00006, respectively) and as PROSITE Accession No. PS00005, PS00007, PS00001, PS00008, PS00006, respectively.

25 In another preferred embodiment, a MP-7 family member is identified further based on the presence of an "N-terminal signal sequence". As used herein, a "signal sequence" includes an amino acid sequence of at least about 15-30, preferably 20-25, more preferably 22 or 23 amino acid residues in length which is located at the extreme N-terminal end of secretory and integral membrane proteins. Such a signal sequence
30 contains large numbers of hydrophobic amino acid residues and is also referred to in the art as a "signal peptide". A signal sequence functions to direct a protein containing such

- 18 -

a sequence to a lipid bilayer. For example, in one embodiment, an MP-7 protein contains a signal sequence containing about amino acids 1-23 of SEQ ID NO:2.

In a preferred embodiment, MP-7 proteins of the invention have an amino acid sequence of about 50-100, more preferably about 100-200, more preferably about 200-300, and even more preferably about 300-350 or 335 amino acid residues in length.

Isolated proteins of the present invention, preferably MP-7 proteins, have an amino acid sequence sufficiently homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 or are encoded by a nucleotide sequence which includes a nucleotide sequence sufficiently homologous to SEQ ID NO:1, SEQ ID NO:3. As used herein, the term "sufficiently homologous" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains have at least about 30-40% homology, preferably 40-50% homology, more preferably 50-60%, and even more preferably 60-70%, 70-80%, or 80-90% homology across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently homologous. Furthermore, amino acid or nucleotide sequences which share at least 30-40%, preferably 40-50%, more preferably 50-60%, 60-70%, 70-80%, or 80-90% homology and share a common functional activity are defined herein as sufficiently homologous.

As used interchangeably herein, a "MP-7 activity", "biological activity of MP-7" or "functional activity of MP-7", refers to an activity exerted by a MP-7 protein, polypeptide or nucleic acid molecule as determined *in vivo*, *in vitro*, or *in situ*, according to standard techniques. In one embodiment, a MP-7 activity is a direct activity, such as an association with a MP-7-target molecule. As used herein, a "target molecule" is a molecule with which a MP-7 protein binds or interacts in nature, such that MP-7-mediated function is achieved. A MP-7 target molecule can be a MP-7 protein or polypeptide of the present invention or a non-MP-7 molecule. For example, a MP-7 target molecule can be a non-MP-7 protein molecule. Alternatively, a MP-7 activity is an indirect activity, such as an activity mediated by

interaction of the MP-7 protein with a MP-7 target molecule such that the target molecule modulates a downstream cellular activity (e.g., interaction of an MP-7 molecule with a MP-7 target molecule can modulate the activity of that target molecule on an immune cell or a cardiac cell).

- 5 In a preferred embodiment, a MP-7 activity is at least one or more of the following activities: (i) interaction of a MP-7 protein with a MP-7 target molecule; (ii) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target molecule is MP-7; (iii) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a ligand, e.g., a mitogenic factor, for example, a phorbol ester; (iv) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is an antibody, e.g., monoclonal antibody, polyclonal antibody; (v) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a monoclonal antibody that interacts with other molecules, e.g., IgG antibodies; (vi) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a receptor, e.g., BCM-1 receptor; (vii) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a receptor, e.g., cell surface receptor mediating cell-cell interaction, for example, cell-cell recognition.

- In yet another preferred embodiment, a MP-7 activity is at least one or more of the following activities: (1) cellular regulation of immune cell types, e.g., T-cells, B-cells, thymocytes and bone marrow progenitor cells, either *in vitro*, *in vivo* or *in situ*; (2) regulation of the cell cycle, e.g., T-cell activation, for example, T-cell differentiation, either *in vitro*, *in vivo* or *in situ*; (3) regulation of the differentiation of multipotent cells, for example, precursor or progenitor cells, e.g., bone marrow progenitor cells; either *in vitro*, *in vivo* or *in situ*; (4) modulation of cell-cell interactions either *in vitro*, *in vivo* or *in situ*; (5) modulation of cardiovascular disorder, e.g., cardiomyopathy, and (6) regulation of differentiation, e.g., cardiovascular differentiation.

- Accordingly, another embodiment of the invention features isolated MP-7 proteins and polypeptides having a MP-7 activity. Preferred proteins are MP-7 proteins having at least one and, preferably two Ig-like domains, preferably at least two extracellular Ig-like domains and, preferably having a MP-7 activity. Additional preferred proteins are MP-7 proteins having a transmembrane domain and, preferably

having a MP-7 activity. Preferred proteins are MP-7 proteins including a cytoplasmic domain and, preferably, a MP-7 activity. In another preferred embodiment, the isolated proteins preferably have a signal sequence. In still another preferred embodiment, the isolated protein is a MP-7 protein having at least one and, preferably, two Ig-like domains, a transmembrane domain, and a cytoplasmic domain and having a MP-7 activity, and preferably, an amino acid sequence sufficiently homologous to an amino acid sequence of SEQ ID NO:2, and optionally a signal sequence and/or propeptide.

MP-7 nucleic acid molecules were identified by screening a cDNA library prepared from a patient suffering from congestive heart failure (described in detail in Example 1). The human MP-7 cDNA, which is approximately 1008 nucleotides in length, encodes a protein which is approximately 335 amino acid residues in length. The human MP-7 protein has at least a first and second extracellular Ig-like domain. A first extracellular Ig-like domain includes, for example, about amino acids 34-107 of SEQ ID NO:2, SEQ ID NO:4. The human MP-7 protein has at least a second extracellular Ig-like domain. A second Ig-like domain includes, for example, about amino acids 144-197 of SEQ ID NO:2, SEQ ID NO:4.

The human MP-7 protein is predicted to be a membrane bound protein which further contains a signal sequence at about amino acids 1-23 of SEQ ID NO:2. The prediction of such a signal peptide can be made, for example, utilizing the computer algorithm SIGNALP (Henrik, *et al.* (1997) *Protein Engineering* 10:1-6). The MP-7 protein wherein the signal peptide has been post-translationally cleaved includes, for example, about 312 amino acids of SEQ ID NO:4. Further, the nucleic acid molecule which encodes such a MP-7 protein without the signal peptide sequence includes, for example, about 936 nucleotides of SEQ ID NO:5. The human MP-7 protein further has at least a transmembrane domain. A transmembrane domain includes, for example, about amino acids 226-250 of SEQ ID NO:2, SEQ ID NO:4. The human MP-7 protein further has at least a cytoplasmic domain. A cytoplasmic domain includes, for example, about amino acids 250-335 of SEQ ID NO:2, SEQ ID NO:4.

The nucleotide sequence of the isolated human MP-7 cDNA and the predicted amino acid sequence of the human MP-7 polypeptide are shown in Figures 1 and 2 and in SEQ ID NOs:1, 3 and 2 respectively. A plasmid containing the full length nucleotide

sequence encoding human MP-7 was deposited with American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on May 20, 1998 and assigned Accession Number 209887. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of
5 Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

Various aspects of the invention are described in further detail in the following subsections:

10

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode MP-7 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify MP-7-encoding nucleic
15 acids (e.g., MP-7 mRNA) and fragments for use as PCR primers for the amplification or mutation of MP-7 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-
20 stranded, but preferably is double-stranded DNA.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the
25 organism from which the nucleic acid is derived. For example, in various embodiments, the isolated MP-7 nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free
30 of other cellular material, or culture medium when produced by recombinant techniques,

- 22 -

or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, the nucleotide
5 sequence of , or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of SEQ ID NO:1, SEQ ID NO:3, the nucleotide sequence of , or the nucleotide sequence of the DNA insert of the plasmid
10 deposited with ATCC as Accession Number 209887, as a hybridization probe, MP-7 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989*).

15 Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert
20 of the plasmid deposited with ATCC as Accession Number 209887.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.
25 Furthermore, oligonucleotides corresponding to MP-7 nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1, corresponds to the human MP-7 cDNA. This cDNA comprises sequences
30 encoding the human MP-7 protein (i.e., "the coding region", from nucleotides 1-1005) and 3' untranslated sequences (nucleotides 1006-1008). Alternatively, the nucleic acid

- 23 -

molecule can comprise only the coding region of SEQ ID NO:1, (e.g., nucleotides 1-1005, corresponding to SEQ ID NO:3).

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide
5 sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as
10 Accession Number 209887, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as
15 Accession Number 209887, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 30-35%, preferably about 35-40%, more preferably at least about 40-45%, more preferably at least about 45-50%, and even more preferably at least about 50-55%, 55-60%, 60-65%,
20 65-70%, 70-75%, 75-80%, 80-85%, 85-90%, or 90-95% or more homologous to the nucleotide sequences shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a portion of any of these nucleotide sequences.

Moreover, the nucleic acid molecule of the invention can comprise only a portion
25 of the nucleic acid sequence (e.g., to the entire length of the nucleotide sequence) of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, for example a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of a MP-7 protein. The nucleotide sequence determined from the cloning of the
30 MP-7 genes allows for the generation of probes and primers designed for use in identifying and/or cloning other MP-7 family members, as well as MP-7 homologues

- 24 -

from other species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 18 or 20, preferably about 22 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, of an anti-sense sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or of a naturally occurring mutant of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887. In an exemplary embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is about 700, preferably 700-800, more preferably 800-900, more preferably 900-1000, and even more preferably 1000-1008 nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887.

Probes based on the MP-7 nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, e.g., the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a MP-7 protein, such as by measuring a level of a MP-7-encoding nucleic acid in a sample of cells from a subject e.g., detecting MP-7 mRNA levels or determining whether a genomic MP-7 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a MP-7 protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, which encodes a polypeptide having a MP-7 biological activity (the biological activities of the MP-7 proteins have previously been described), expressing the encoded portion of the MP-7 protein (e.g., by

- 25 -

recombinant expression *in vitro*) and assessing the activity of the encoded portion of the MP-7 protein:

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence
5 of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, due to degeneracy of the genetic code and thus encode the same MP-7 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887. In another embodiment, an isolated nucleic acid molecule
10 of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:2, SEQ ID NO:4.

In addition to the MP-7 nucleotide sequences shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, it will be appreciated by those skilled in the art
15 that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the MP-7 proteins may exist within a population (e.g., the human population). Such genetic polymorphism in the MP-7 genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame
20 encoding a MP-7 protein, preferably a mammalian MP-7 protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a MP-7 gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in MP-7 genes that are the result of natural allelic variation and that do not alter the functional activity of a MP-7 protein are intended to be within the scope of the
25 invention.

Allelic variants of human MP-7 include both functional and non-functional MP-7 proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human MP-7 protein that maintain the ability to bind an MP-7 ligand and/or modulate cardiomyocyte function. Functional allelic variants will typically
30 contain only conservative substitution of one or more amino acids of SEQ ID NO:2 or

substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human MP-7 protein that do not have the ability to either bind an MP-7
5 ligand and/or modulate cardiomyocyte function. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2 or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues of the human
10 MP-7 protein. Orthologues of the human MP-7 protein are proteins that are isolated from non-human organisms and possess the same MP-7 ligand binding and/or modulation of cardiomyocyte function capabilities of the human MP-7 protein. Orthologues of the human MP-7 protein can readily be identified as comprising an amino acid sequence that is substantially homologous to SEQ ID NO:2.

15 Moreover, nucleic acid molecules encoding other MP-7 family members (e.g., MP-7-2), and thus which have a nucleotide sequence which differs from the MP-7 sequences of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887 are intended to be within the scope of the invention. For example, a MP-7-2 cDNA can be identified
20 based on the nucleotide sequence of human MP-7. Moreover, nucleic acid molecules encoding MP-7 proteins from different species, and thus which have a nucleotide sequence which differs from the MP-7 sequences of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887 are intended to be within the scope of the invention. For
25 example, an mouse MP-7 cDNA can be identified based on the nucleotide sequence of a human MP-7.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the MP-7 cDNAs of the invention can be isolated based on their homology to the MP-7 nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof,
30 as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

- 27 -

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with

5 ATCC as Accession Number 209887. In other embodiment, the nucleic acid is at least 30, 50, 100, 250 or 500 nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least about 25%, 30%, 35%, 40%, 42%, 43%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more

10 homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 42%, 43%, more preferably at least about 50%, more preferably at least about 60%, even more preferably at least about 70%, more preferably at least about 80%, even more preferably at least about 85%, even more preferably at least about 85% or 90% homologous to each other typically remain

15 hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an

20 isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, SEQ ID NO:3 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

25 In addition to naturally-occurring allelic variants of the MP-7 sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, thereby leading to changes in the amino acid sequence of

30 the encoded MP-7 proteins, without altering the functional ability of the MP-7 proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-

- 28 -

essential" amino acid residues can be made in the sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of MP-7 (e.g., the sequence of SEQ ID NO:2, SEQ ID NO:4) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the MP-7 proteins of the present invention, are predicted to be particularly unamenable to alteration (e.g., the ten conserved cysteines involved in forming disulfide linkages or the conserved histidine, aspartate, or serine of the active enzymatic site). Moreover, amino acid residues that are defined by the MP-7 Ig-like domains, transmembrane domain, and cytoplasmic domain are particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the MP-7 proteins of the present invention and other members of the Ly-9 and LY-6 superfamily or protein families containing GPI-anchored cell surface lymphocyte specific alloantigens are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding MP-7 proteins that contain changes in amino acid residues that are not essential for activity. Such MP-7 proteins differ in amino acid sequence from SEQ ID NO:2, SEQ ID NO:4 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 20%, 25%, 30%, 35%, 40%, 42%, 43%, 45%, 48%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or more homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4. Preferably, the protein encoded by the nucleic acid molecule is at least about 65-70% homologous to SEQ ID NO:2, SEQ ID NO:4, more preferably at least about 75-80% homologous to SEQ ID NO:2, SEQ ID NO:4, even more preferably at least about 85-90% homologous to SEQ ID NO:2, SEQ ID NO:4, and most preferably at least about 95% homologous to SEQ ID NO:2, SEQ ID NO:4 (e.g., the entire amino acid sequence of SEQ ID NO:2, SEQ ID NO:4).

- 29 -

An isolated nucleic acid molecule encoding a MP-7 protein homologous to the protein of SEQ ID NO:2, SEQ ID NO:4 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis.

Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a MP-7 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a MP-7 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for MP-7 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In a preferred embodiment, a mutant MP-7 protein can be assayed for the ability to (1) modulate cellular regulation of immune cell types, e.g., T-cells, B-cells, thymocytes and bone marrow progenitor cells, either *in vitro*, *in vivo* or *in situ*; (2)

- 30 -

modulate the regulation of the cell cycle, e.g., T-cell activation, for example, T-cell differentiation, either *in vitro*, *in vivo* or *in situ*; (3) modulate the regulation of the differentiation of multipotent cells, for example, precursor or progenitor cells, e.g., bone marrow progenitor cells; either *in vitro*, *in vivo* or *in situ*; (4) modulate the regulation of cell-cell interactions either *in vitro*, *in vivo* or *in situ*; (5) modulate of cardiovascular disorder, e.g., cardiomyopathy, and (6) regulate of differentiation, e.g., cardiovascular differentiation. In addition to the nucleic acid molecules encoding MP-7 proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire MP-7 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding MP-7. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (e.g., the coding region of human MP-7 corresponds to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding MP-7. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding MP-7 disclosed herei, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of MP-7 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of MP-7 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of MP-7 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An

- 31 -

antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to

5 increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-

10 acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-

15 methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-

20 carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

25 The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a MP-7 protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the

30 case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of

administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to
5 receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is
10 placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-*o*-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).
15

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are
20 capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave MP-7 mRNA transcripts to thereby inhibit translation of MP-7 mRNA. A ribozyme having specificity for a MP-7-encoding nucleic acid can be
25 designed based upon the nucleotide sequence of a MP-7 cDNA disclosed herein (i.e., SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number 209887). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in
30 a MP-7-encoding mRNA. See, e.g., Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, MP-7 mRNA can be used to select a

catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, MP-7 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the MP-7 (e.g., the MP-7 promoter and/or enhancers) to form triple helical structures that prevent transcription of the MP-7 gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the MP-7 nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* PNAS 93: 14670-675.

PNAs of MP-7 nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of MP-7 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

- 34 -

In another embodiment, PNAs of MP-7 can be modified, (e.g., to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of MP-7 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (e.g., RNase H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. US.* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134, published April 25, 1988). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, e.g., Krol *et al.* (1988) *BioTechniques* 6:958-976) or intercalating agents. (See, e.g., Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

II. Isolated MP-7 Proteins and Anti-MP-7 Antibodies

One aspect of the invention pertains to isolated MP-7 proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-MP-7 antibodies. In one embodiment, native MP-7 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, MP-7 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a MP-7 protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the MP-7 protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of MP-7 protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of MP-7 protein having less than about 30% (by dry weight) of non-MP-7 protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-MP-7 protein, still more preferably less than about 10% of non-MP-7 protein, and most preferably less than about 5% non-MP-7 protein. When the MP-7 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of MP-7 protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of MP-7 protein having less than about 30% (by dry weight) of chemical precursors or non-MP-7 chemicals, more preferably less than about 20% chemical precursors or non-MP-7 chemicals, still more preferably less than about 10%

- 36 -

chemical precursors or non-MP-7 chemicals, and most preferably less than about 5% chemical precursors or non-MP-7 chemicals.

Biologically active portions of a MP-7 protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid
5 sequence of the MP-7 protein, e.g., the amino acid sequence shown in SEQ ID NO:2, SEQ ID NO:4, which include less amino acids than the full length MP-7 proteins, and exhibit at least one activity of a MP-7 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the MP-7 protein. A biologically active portion of a MP-7 protein can be a polypeptide which is, for example,
10 10, 25, 50, 100 or more amino acids in length.

In one embodiment, a biologically active portion of a MP-7 protein comprises at least an Ig-like domain. In another embodiment, a biologically active portion of a MP-7 protein comprises at least a first or second or both a first and second Ig-like domains. In another embodiment, a biologically active portion of a MP-7 protein comprises at least a
15 transmembrane domain. In another embodiment, a biologically active portion of a MP-7 protein comprises at least a cytoplasmic domain. In another embodiment, a biologically active portion of a MP-7 protein comprises at least an Ig-like domain or at least a first and second Ig-like domain and/or a transmembrane domain and/or a cytoplasmic domain.

20 It is to be understood that a preferred biologically active portion of a MP-7 protein of the present invention may contain at least one of the above-identified structural domains. A more preferred biologically active portion of a MP-7 protein may contain at least two of the above-identified structural domains. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be
25 prepared by recombinant techniques and evaluated for one or more of the functional activities of a native MP-7 protein.

In a preferred embodiment, the MP-7 protein has an amino acid sequence shown in SEQ ID NO:2, SEQ ID NO:4. In other embodiments, the MP-7 protein is substantially homologous to SEQ ID NO:2, SEQ ID NO:4, and retains the functional
30 activity of the protein of SEQ ID NO:2, SEQ ID NO:4, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in

subsection I above. Accordingly, in another embodiment, the MP-7 protein is a protein which comprises an amino acid sequence at least about 31%, 32% homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 and retains the functional activity of the MP-7 proteins of SEQ ID NO:2, SEQ ID NO:4, respectively. Preferably, the protein is at least about 30-35% homologous to SEQ ID NO:2, SEQ ID NO:4, more preferably at least about 35-40% homologous to SEQ ID NO:2, SEQ ID NO:4, even more preferably at least about 40-45% homologous to SEQ ID NO:2, SEQ ID NO:4, and even more preferably at least about 45-50%, 50-55%, 55-60%, 60-65%, 65-70%, 70-75%, 75-80%, 80-85%, 85-90%, or 90-95% or more homologous to SEQ ID NO:2, SEQ ID NO:4.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (e.g., when aligning a second sequence to the MP-7 amino acid sequence of SEQ ID NO:2 having 167 amino acid residues, at least 100, preferably at least 150, more preferably at least 200, and even more preferably at least 250, 300 or 335 amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

- 38 -

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has
5 been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at
10 <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989) which has been incorporated into the ALIGN program (version 2.0) (available at
15 <http://vega.igh.cnrs.fr/bin/align-guess.cgi>), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed
20 using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to MP-7 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain
25 amino acid sequences homologous to MP-7 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

The invention also provides MP-7 chimeric or fusion proteins. As used herein, a MP-7 "chimeric protein" or "fusion protein" comprises a MP-7 polypeptide operatively linked to a non-MP-7 polypeptide. A "MP-7 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to MP-7, whereas a "non-MP-7 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the MP-7 protein, e.g., a protein which is different from the MP-7 protein and which is derived from the same or a different organism. Within a MP-7 fusion protein the MP-7 polypeptide can correspond to all or a portion of a MP-7 protein. In a preferred embodiment, a MP-7 fusion protein comprises at least one biologically active portion of a MP-7 protein. In another preferred embodiment, a MP-7 fusion protein comprises at least two biologically active portions of a MP-7 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the MP-7 polypeptide and the non-MP-7 polypeptide are fused in-frame to each other. The non-MP-7 polypeptide can be fused to the N-terminus or C-terminus of the MP-7 polypeptide.

For example, in one embodiment, the fusion protein is a GST-MP-7 fusion protein in which the MP-7 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant MP-7.

In another embodiment, the fusion protein is a MP-7 protein containing a heterologous signal sequence at its N-terminus. For example, the native MP-7 signal sequence (i.e., about amino acids 1 to 23 of SEQ ID NO:2) can be removed and replaced with a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of MP-7 can be increased through use of a heterologous signal sequence.

The MP-7 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The MP-7 fusion proteins can be used to affect the bioavailability of a MP-7 target molecule. Use of MP-7 fusion proteins may be useful therapeutically for the treatment of cardiovascular disorders (e.g., congestive heart failure). Moreover, the MP-7-fusion proteins of the invention can be used as immunogens to produce anti-MP-7 antibodies in a subject, to

- 40 -

purify MP-7 ligands and in screening assays to identify molecules which inhibit the interaction of MP-7 with a MP-7 target molecule.

Preferably, a MP-7 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A MP-7-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the MP-7 protein.

The present invention also pertains to variants of the MP-7 proteins which function as either MP-7 agonists (mimetics) or as MP-7 antagonists. Variants of the MP-7 proteins can be generated by mutagenesis, e.g., discrete point mutation or truncation of a MP-7 protein. An agonist of the MP-7 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a MP-7 protein. An antagonist of a MP-7 protein can inhibit one or more of the activities of the naturally occurring form of the MP-7 protein by, for example, competitively inhibiting the protease activity of a MP-7 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the MP-7 protein.

In one embodiment, variants of a MP-7 protein which function as either MP-7 agonists (mimetics) or as MP-7 antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of a MP-7 protein for MP-7 protein agonist or antagonist activity. In one embodiment, a variegated library of MP-7 variants is
5 generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of MP-7 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential MP-7 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage
10 display) containing the set of MP-7 sequences therein. There are a variety of methods which can be used to produce libraries of potential MP-7 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision,
15 in one mixture, of all of the sequences encoding the desired set of potential MP-7 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

20 In addition, libraries of fragments of a MP-7 protein coding sequence can be used to generate a variegated population of MP-7 fragments for screening and subsequent selection of variants of a MP-7 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR
25 fragment of a MP-7 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal,
30 C-terminal and internal fragments of various sizes of the MP-7 protein.

- 42 -

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of MP-7 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify MP-7 variants (Arkin and Yourvan (1992) *PNAS* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated MP-7 library. For example, a library of expression vectors can be transfected into a cell line which ordinarily synthesizes and secretes MP-7. The transfected cells are then cultured such that MP-7 and a particular mutant MP-7 are secreted and the effect of expression of the mutant on MP-7 activity in cell supernatants can be detected, e.g., by any of a number of enzymatic assays. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of MP-7 activity, and the individual clones further characterized.

An isolated MP-7 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind MP-7 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length MP-7 protein can be used or, alternatively, the invention provides antigenic peptide fragments of MP-7 for use as immunogens. The antigenic peptide of MP-7 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2, SEQ ID NO:4 and encompasses an epitope of MP-7 such that an antibody raised against the peptide forms a specific immune complex with MP-7. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more

- 43 -

preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of MP-7 that are located on the surface of the protein, e.g., hydrophilic regions.

5 A MP-7 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed MP-7 protein or a chemically synthesized MP-7 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or
10 similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic MP-7 preparation induces a polyclonal anti-MP-7 antibody response.

Accordingly, another aspect of the invention pertains to anti-MP-7 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that
15 contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as MP-7. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')₂ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind MP-7. The term "monoclonal antibody" or
20 "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of MP-7. A monoclonal antibody composition thus typically displays a single binding affinity for a particular MP-7 protein with which it immunoreacts.

25 Polyclonal anti-MP-7 antibodies can be prepared as described above by immunizing a suitable subject with a MP-7 immunogen. The anti-MP-7 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized MP-7. If desired, the antibody molecules directed against MP-7 can be isolated from the mammal (e.g.,
30 from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization,

- 44 -

e.g., when the anti-MP-7 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *PNAS* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a MP-7 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds MP-7.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-MP-7 monoclonal antibody (see, e.g., G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These

myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days
5 because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind MP-7, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-MP-7 antibody can be identified and isolated by screening a
10 recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with MP-7 to thereby isolate immunoglobulin library members that bind MP-7. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally,
15 examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication
20 No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281;
25 Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *PNAS* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *PNAS* 88:7978-7982; and McCafferty *et al.* *Nature* (1990) 348:552-554.

Additionally, recombinant anti-MP-7 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by

5 recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567;

10 Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.*

15 (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-MP-7 antibody (e.g., monoclonal antibody) can be used to isolate MP-7 by standard techniques, such as affinity chromatography or immunoprecipitation. An

20 anti-MP-7 antibody can facilitate the purification of natural MP-7 from cells and of recombinantly produced MP-7 expressed in host cells. Moreover, an anti-MP-7 antibody can be used to detect MP-7 protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the MP-7 protein. Anti-MP-7 antibodies can be used diagnostically to monitor protein levels in tissue as part of

25 a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes

30 include horseradish peroxidase, alkaline phosphatase, -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include

- 47 -

streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include
5 luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression
10 vectors, containing a nucleic acid encoding a MP-7 protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional
15 DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host
20 genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector.
25 However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which
30 means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is

- 48 -

operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., MP-7 proteins, mutant forms of MP-7 proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of MP-7 proteins in prokaryotic or eukaryotic cells. For example, MP-7 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein;

2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in MP-7 activity assays, (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific for MP-7 proteins, for example. In a preferred embodiment, a MP-7 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another

strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard
5 DNA synthesis techniques.

In another embodiment, the MP-7 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen
10 Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, MP-7 proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

15 In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements.
20 For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,
25 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable
30 tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988)

Adv. Immunol. 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477),
5 pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989)
10 *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an
15 RNA molecule which is antisense to MP-7 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The
20 antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a
25 molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny
30 of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be

identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a MP-7 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or
5 mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized
10 techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al. (Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor*
15 *Laboratory Press, Cold Spring Harbor, NY, 1989)*, and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is
20 generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a MP-7 protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic
25 acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) a MP-7 protein. Accordingly, the invention further provides methods for producing a MP-7 protein using the host cells of
30 the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding a MP-7 protein has

been introduced) in a suitable medium such that a MP-7 protein is produced. In another embodiment, the method further comprises isolating a MP-7 protein from the medium or the host cell.

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which MP-7-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous MP-7 sequences have been introduced into their genome or homologous recombinant animals in which endogenous MP-7 sequences have been altered. Such animals are useful for studying the function and/or activity of a MP-7 and for identifying and/or evaluating modulators of MP-7 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous MP-7 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a MP-7-encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The MP-7 cDNA sequence of SEQ ID NO:1, SEQ ID NO:3 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human MP-7 gene, such as a mouse or rat MP-7 gene, can be used as a transgene. Alternatively, a MP-7 gene homologue, such as a MP-7-2 gene can be isolated based on hybridization to the MP-7 cDNA sequences of

SEQ ID NO:1, SEQ ID NO:3, or the DNA insert of the plasmid deposited with ATCC as Accession Number 209887 (described further in subsection I above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a MP-7 transgene to direct expression of a MP-7 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a MP-7 transgene in its genome and/or expression of MP-7 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a MP-7 protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a MP-7 gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the MP-7 gene. The MP-7 gene can be a human gene (e.g., the cDNA of), but more preferably, is a non-human homologue of a human MP-7 gene (e.g., a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3). For example, a mouse MP-7 gene can be used to construct a homologous recombination vector suitable for altering an endogenous MP-7 gene in the mouse genome. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous MP-7 gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous MP-7 gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous MP-7 protein). In the homologous

- 55 -

recombination vector, the altered portion of the MP-7 gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the MP-7 gene to allow for homologous recombination to occur between the exogenous MP-7 gene carried by the vector and an endogenous MP-7 gene in an embryonic stem cell. The additional flanking MP-7

5 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in

10 which the introduced MP-7 gene has homologously recombined with the endogenous MP-7 gene are selected (see e.g., Li, E. *et al.* (1992) *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see e.g., Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo

15 can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant

20 animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-humans animals can be produced which

25 contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If

30 a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are

- 56 -

required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be
5 produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. Alternatively, a cell, e.g., an embryonic stem cell, from the inner cell mass of a developing embryo can be transformed with a preferred transgene. Alternatively, a cell, e.g., a somatic cell, from
10 cell culture line can be transformed with a preferred transgene and induced to exit the growth cycle and enter G₀ phase. The cell can then be fused, e.g., through the use of electrical pulses, to an enucleated mammalian oocyte. The reconstructed oocyte is then cultured such that it develops to a morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal
15 will be a clone of the animal from which the nuclear donor cell, e.g., the somatic cell, is isolated.

IV. Pharmaceutical Compositions

The MP-7 nucleic acid molecules, MP-7 proteins, and anti-MP-7 antibodies (also
20 referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and
25 antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the
30 compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions
5 used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic
10 acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

15 Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In
20 all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid
25 polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol,
30 ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium

chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active
5 compound (e.g., a MP-7 protein or anti-MP-7 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In
10 the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They
15 can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically
20 compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant
25 such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant,
30 e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and
5 therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

10 The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the
15 method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine
20 useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by
25 stereotactic injection (see *e.g.*, Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the
30 pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

5 The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (e.g., therapeutic and prophylactic).

10 As described herein, a MP-7 protein of the invention has one or more of the following activities: (i) interaction of a MP-7 protein with a MP-7 target molecule; (ii) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target molecule is MP-7; (iii) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a ligand, e.g., a mitogenic factor, for example, a phorbol
15 ester; (iv) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is an antibody, e.g., monoclonal antibody, polyclonal antibody; (v) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a monoclonal antibody that interacts with other molecules, e.g., IgG antibodies; (vi) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a receptor, e.g.,
20 BCM-1 receptor; (vii) interaction of a MP-7 protein with a MP-7 target molecule, wherein the MP-7 target is a receptor, e.g., cell surface receptor mediating cell-cell interaction, for example, cell-cell recognition.

 Further as described herein, a MP-7 protein of the invention has one or more of the above activities and can thus be used in, for example, the: (1) cellular regulation of
25 immune cell types, e.g., T-cells, B-cells, thymocytes and bone marrow progenitor cells, either *in vitro*, *in vivo* or *in situ*; (2) regulation of the cell cycle, e.g., T-cell activation, for example, T-cell differentiation, either *in vitro*, *in vivo* or *in situ*; (3) regulation of the differentiation of multipotent cells, for example, precursor or progenitor cells, e.g., bone marrow progenitor cells; either *in vitro*, *in vivo* or *in situ*; (4) modulation of cell-cell
30 interactions either *in vitro*, *in vivo* or *in situ*; (5) modulation of cardiovascular disorder,

e.g., cardiomyopathy, and (6) regulation of differentiation, e.g., cardiovascular differentiation.

The isolated nucleic acid molecules of the invention can be used, for example, to express MP-7 protein (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect MP-7 mRNA (e.g., in a biological sample) or a genetic alteration in a MP-7 gene, and to modulate MP-7 activity, as described further below. The MP-7 proteins can be used to treat disorders characterized by insufficient or excessive production of a MP-7 or MP-7 target molecules. In addition, the MP-7 proteins can be used to screen for naturally occurring MP-7 target molecules, to screen for drugs or compounds which modulate MP-7 activity, as well as to treat disorders characterized by insufficient or excessive production of MP-7 protein or production of MP-7 protein forms which have decreased or aberrant activity compared to MP-7 wild type protein. Moreover, the anti-MP-7 antibodies of the invention can be used to detect and isolate MP-7 proteins, regulate the bioavailability of MP-7 proteins, and modulate MP-7 activity.

Accordingly one embodiment of the present invention involves a method of use (e.g., a diagnostic assay, prognostic assay, or a prophylactic/therapeutic method of treatment) wherein a molecule of the present invention (e.g., a MP-7 protein, MP-7 nucleic acid, or a MP-7 modulator) is used, for example, to diagnose, prognose and/or treat a disease and/or condition in which any of the aforementioned activities (i.e., activities (i) - (vii) and (1) - (6) in the above paragraph) is indicated. In another embodiment, the present invention involves a method of use (e.g., a diagnostic assay, prognostic assay, or a prophylactic/therapeutic method of treatment) wherein a molecule of the present invention (e.g., a MP-7 protein, MP-7 nucleic acid, or a MP-7 modulator) is used, for example, for the diagnosis, prognosis, and/or treatment of subjects, preferably a human subject, in which any of the aforementioned activities is pathologically perturbed. In a preferred embodiment, the methods of use (e.g., diagnostic assays, prognostic assays, or prophylactic/therapeutic methods of treatment) involve administering to a subject, preferably a human subject, a molecule of the present invention (e.g., a MP-7 protein, MP-7 nucleic acid, or a MP-7 modulator) for the diagnosis, prognosis, and/or therapeutic treatment. In another embodiment, the methods

of use (e.g., diagnostic assays, prognostic assays, or prophylactic/therapeutic methods of treatment) involve administering to a human subject a molecule of the present invention (e.g., a MP-7 protein, MP-7 nucleic acid, or a MP-7 modulator).

5 A. Screening Assays:

 The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which bind to MP-7 proteins, have a stimulatory or inhibitory effect on, for example, MP-7 expression or MP-7 activity, or
10 have a stimulatory or inhibitory effect on, for example, the activity of an MP-7 target molecule.

 In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of a MP-7 protein or polypeptide or biologically active portion thereof. In another
15 embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a MP-7 protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution
20 phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

25 Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in
30 Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage
5 (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses membrane-bound form of a MP-7 protein or biologically active portion
10 thereof on the cell surface is contacted with a test compound and the ability of the test compound to modulate MP-7 activity determined. Determining the ability of the test compound to modulate MP-7 activity can be accomplished by monitoring the bioactivity of the MP-7 protein or biologically active portion thereof. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to
15 modulate MP-7 activity can be accomplished, for example, by coupling the test compound, the MP-7 protein or biologically active portion thereof with a radioisotope or enzymatic label such that binding of the test compound, the MP-7 protein or biologically active portion thereof to its cognate target molecule can be determined by detecting the labeled test compound, MP-7 protein or biologically active portion thereof in a complex.
20 For example, test compounds (e.g., MP-7 protein or biologically active portion thereof) can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label
25 detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (e.g., MP-7 protein or biologically active portion thereof) to interact with its cognate target molecule without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with its cognate
30 target molecule without the labeling of either the compound or the receptor. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer"

(e.g., Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between compound and receptor.

5 In a preferred embodiment, the assay comprises contacting a cell which expresses membrane-bound form of MP-7 protein, or a biologically active portion thereof, on the cell surface with a known compound which binds MP-7 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to modulate the activity of the MP-7 protein or biologically active
10 portion thereof, wherein determining the ability of the test compound to modulate the activity of the MP-7 protein or biologically active portion thereof, comprises determining the ability of the test compound to modulate a biological activity of the MP-7 expressing cell (e.g., determining the ability of the test compound to modulate cardiac cell function, immune cell activation, cell-cell interactions, and/or protein:protein
15 interactions).

 In another preferred embodiment, the assay comprises contacting a cell which is responsive to a MP-7 protein or biologically active portion thereof, with a MP-7 protein or biologically-active portion thereof, to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to
20 modulate the activity of the MP-7 protein or biologically active portion thereof, wherein determining the ability of the test compound to modulate the activity of the MP-7 protein or biologically active portion thereof comprises determining the ability of the test compound to modulate a biological activity of the MP-7-responsive cell (e.g., determining the ability of the test compound to modulate cardiac cell function, immune
25 cell activation, cell-cell interactions, and/or protein:protein interactions).

 In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of MP-7 protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (e.g. stimulate or inhibit) the activity of the MP-7 protein
30 or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of MP-7 or a biologically active portion thereof can be

- 66 -

accomplished, for example, by determining the ability of the MP-7 protein to bind to or interact with a MP-7 target molecule. As used herein, a "target molecule" is a molecule with which a MP-7 protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a MP-7 protein, a molecule on the surface of a second
5 cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A MP-7 target molecule can be a non-MP-7 molecule or a MP-7 protein or polypeptide of the present invention. In one embodiment, a MP-7 target molecule is a component of a signal transduction pathway which facilitates transduction of an extracellular signal (e.g. a signal generated by
10 binding of a compound to a membrane-bound MP-7 molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein which has catalytic activity or a protein which facilitates the association of downstream signaling molecules with MP-7. Alternatively, the target molecule can be a substrate for a catalytic activity of the MP-7 protein.

15 Determining the ability of the MP-7 protein to bind to or interact with a MP-7 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the MP-7 protein to bind to or interact with a MP-7 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target
20 molecule can be determined by detecting induction of a cellular second messenger of the target (i.e. intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a target-regulated cellular
25 response, for example, cardiac cell function, immune cell activation, cell-cell interactions, and/or protein:protein interactions.

In yet another embodiment, an assay of the present invention is a cell-free assay in which a MP-7 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the MP-7 protein or
30 biologically active portion thereof is determined. Binding of the test compound to the MP-7 protein can be determined either directly or indirectly as described above. In a

- 67 -

preferred embodiment, the assay includes contacting the MP-7 protein or biologically active portion thereof with a known compound which binds MP-7 (e.g., a MP-7 target molecule) to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a MP-7 protein,

- 5 wherein determining the ability of the test compound to interact with a MP-7 protein comprises determining the ability of the test compound to preferentially bind to MP-7 or biologically active portion thereof as compared to the known compound.

- In another embodiment, the assay is a cell-free assay in which a MP-7 protein or biologically active portion thereof is contacted with a test compound and the ability of
- 10 the test compound to modulate (e.g., stimulate or inhibit) the activity of the MP-7 protein or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a MP-7 protein can be accomplished, for example, by determining the ability of the MP-7 protein to bind to a MP-7 target molecule by one of the methods described above for determining direct binding.
- 15 Determining the ability of the MP-7 protein to bind to a MP-7 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants
- 20 (e.g., BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules. In an alternative embodiment, determining the ability of the test compound to modulate the activity of MP-7 can be accomplished by determining the ability of the MP-7 protein further modulate a MP-7 target molecule. For example, the catalytic/enzymatic activity
- 25 of the target molecule on an appropriate substrate can be determined as previously described.

- In an alternative embodiment, determining the ability of the test compound to modulate the activity of a MP-7 protein can be accomplished by determining the ability of the MP-7 protein to further modulate the activity of a downstream effector of a MP-7
- 30 target molecule. For example, the activity of the effector molecule on an appropriate

- 68 -

target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a MP-7 protein or biologically active portion thereof with a known compound which binds the MP-7 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the MP-7 protein, wherein determining the ability of the test compound to interact with the MP-7 protein comprises determining the ability of the MP-7 protein to preferentially bind to or modulate the activity of a MP-7 target molecule.

The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of isolated proteins (e.g. MP-7 proteins or biologically active portions thereof or receptors to which MP-7 targets bind). In the case of cell-free assays in which a membrane-bound form of an isolated protein is used (e.g., a cell surface receptor) it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the isolated protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either MP-7 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a MP-7 protein, or interaction of a MP-7 protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ MP-7 fusion proteins or glutathione-S-

transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or MP-7 protein, and the mixture incubated under conditions conducive to
5 complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of MP-7 binding or activity determined
10 using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a MP-7 protein or a MP-7 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated MP-7 protein or target molecules can be prepared from biotin-NHS (N-
15 hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with MP-7 protein or target molecules but which do not interfere with binding of the MP-7 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or MP-7
20 protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the MP-7 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the MP-7 protein or target molecule.

25 In another embodiment, modulators of MP-7 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of MP-7 mRNA or protein in the cell is determined. The level of expression of MP-7 mRNA or protein in the presence of the candidate compound is compared to the level of expression of MP-7 mRNA or protein in the absence of the candidate compound. The
30 candidate compound can then be identified as a modulator of MP-7 expression based on this comparison. For example, when expression of MP-7 mRNA or protein is greater

- 70 -

(statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of MP-7 mRNA or protein expression. Alternatively, when expression of MP-7 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of MP-7 mRNA or protein expression. The level of MP-7 mRNA or protein expression in the cells can be determined by methods described herein for detecting MP-7 mRNA or protein.

In yet another aspect of the invention, the MP-7 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with MP-7 ("MP-7-binding proteins" or "MP-7-bp") and are involved in MP-7 activity. Such MP-7-binding proteins are also likely to be involved in the propagation of signals by the MP-7 proteins or MP-7 targets as, for example, downstream elements of a MP-7-mediated signaling pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a MP-7 protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a MP-7-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the MP-7 protein.

- 71 -

This invention further pertains to novel agents identified by the above-described screening assays and to processes for producing such agents by use of these assays. Accordingly, in one embodiment, the present invention includes a compound or agent obtainable by a method comprising the steps of any one of the aforementioned screening assays (e.g., cell-based assays or cell-free assays). For example, in one embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a cell which expresses a MP-7 target molecule with a test compound and the determining the ability of the test compound to bind to, or modulate the activity of, the MP-7 target molecule. In another embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a cell which expresses a MP-7 target molecule with a MP-7 protein or biologically-active portion thereof, to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with, or modulate the activity of, the MP-7 target molecule. In another embodiment, the invention includes a compound or agent obtainable by a method comprising contacting a MP-7 protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to, or modulate (e.g., stimulate or inhibit) the activity of, the MP-7 protein or biologically active portion thereof. In yet another embodiment, the present invention included a compound or agent obtainable by a method comprising contacting a MP-7 protein or biologically active portion thereof with a known compound which binds the MP-7 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with, or modulate the activity of the MP-7 protein.

Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a MP-7 modulating agent, an antisense MP-7 nucleic acid molecule, a MP-7-specific antibody, or a MP-7-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this

- 72 -

invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

The present invention also pertains to uses of novel agents identified by the above-described screening assays for diagnoses, prognoses, and treatments as described herein. Accordingly, it is within the scope of the present invention to use such agents in the design, formulation, synthesis, manufacture, and/or production of a drug or pharmaceutical composition for use in diagnosis, prognosis, or treatment, as described herein. For example, in one embodiment, the present invention includes a method of synthesizing or producing a drug or pharmaceutical composition by reference to the structure and/or properties of a compound obtainable by one of the above-described screening assays. For example, a drug or pharmaceutical composition can be synthesized based on the structure and/or properties of a compound obtained by a method in which a cell which expresses a MP-7 target molecule is contacted with a test compound and the ability of the test compound to bind to, or modulate the activity of, the MP-7 target molecule is determined. In another exemplary embodiment, the present invention includes a method of synthesizing or producing a drug or pharmaceutical composition based on the structure and/or properties of a compound obtainable by a method in which a MP-7 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to, or modulate (e.g., stimulate or inhibit) the activity of, the MP-7 protein or biologically active portion thereof is determined.

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the MP-7

5 nucleotide sequences, described herein, can be used to map the location of the MP-7 genes on a chromosome. The mapping of the MP-7 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, MP-7 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the MP-7 nucleotide sequences. Computer analysis
10 of the MP-7 sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the MP-7 sequences will yield an amplified fragment.

15 Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains
20 the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell
25 hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the MP-7 nucleotide sequences to
30 design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be

used to map a 90, 1p, or 1v sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *PNAS*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

5 Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A
10 pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will
15 suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for
20 marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the
25 physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes),
30 described in, for example, Egeland, J. *et al.* (1987) *Nature*, 325:783-787.

- 75 -

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the MP-7 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease.

- 5 Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

10

2. Tissue Typing

- The MP-7 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for
- 15 identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers
- 20 for RFLP (described in U.S. Patent 5,272,057).

- Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the MP-7 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the
- 25 sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

- Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the
- 30 present invention can be used to obtain such identification sequences from individuals and from tissue. The MP-7 nucleotide sequences of the invention uniquely represent

- 76 -

portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1, SEQ ID NO:3, can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from MP-7 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of Partial MP-7 Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular

individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1, SEQ ID NO:3 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the MP-7 nucleotide sequences or portions thereof, e.g., fragments derived from the noncoding regions of SEQ ID NO:1, SEQ ID NO:3, having a length of at least 20 bases, preferably at least 30 bases.

The MP-7 nucleotide sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such MP-7 probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, e.g., MP-7 primers or probes can be used to screen tissue culture for contamination (i.e. screen for the presence of a mixture of different types of cells in a culture).

C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining MP-7 protein and/or nucleic acid expression as well as MP-7 activity, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant MP-7 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with MP-7 protein, nucleic acid expression or activity. For example, mutations in a MP-7 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically

treat an individual prior to the onset of a disorder characterized by or associated with MP-7 protein, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of MP-7 in clinical trials.

5 These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of MP-7 protein or nucleic acid in a biological sample involves obtaining a biological sample from a test
10 subject and contacting the biological sample with a compound or an agent capable of detecting MP-7 protein or nucleic acid (e.g., mRNA, genomic DNA) that encodes MP-7 protein such that the presence of MP-7 protein or nucleic acid is detected in the biological sample. A preferred agent for detecting MP-7 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to MP-7 mRNA or genomic DNA.
15 The nucleic acid probe can be, for example, a full-length MP-7 nucleic acid, such as the nucleic acid of SEQ ID NO: 1 (or that of , or the DNA insert of the plasmid deposited with ATCC as Accession Number 209887, or a portion thereof), such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to MP-7 mRNA or
20 genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting MP-7 protein is an antibody capable of binding to MP-7 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof
25 (e.g., Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody
30 using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term

"biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect MP-7 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of MP-7 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of MP-7 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of MP-7 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of MP-7 protein include introducing into a subject a labeled anti-MP-7 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting MP-7 protein, mRNA, or genomic DNA, such that the presence of MP-7 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of MP-7 protein, mRNA or genomic DNA in the control sample with the presence of MP-7 protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of MP-7 in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting MP-7 protein or mRNA in a biological sample; means for determining the amount of MP-7 in the sample; and means for comparing the amount of MP-7 in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect MP-7 protein or nucleic acid.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant MP-7 expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with MP-7 protein, nucleic acid expression or activity such a cardiovascular disorder (e.g., cardiomyopathy).

Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a differentiative disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant MP-7 expression or activity in which a test sample is obtained from a subject and MP-7 protein or nucleic acid (e.g., mRNA, genomic DNA) is detected, wherein the presence of MP-7 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant MP-7 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant MP-7 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for cardiovascular disorder (e.g., cardiomyopathy). For example, such methods can be used to determine whether a subject can be effectively treated with an agent for cardiovascular disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant MP-7 expression or activity in which a test sample is obtained and MP-7 protein or nucleic acid expression or activity is detected (e.g., wherein the abundance of MP-7 protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant MP-7 expression or activity.)

The methods of the invention can also be used to detect genetic alterations in a MP-7 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by aberrant cellular function. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a MP-7-protein, or the mis-expression of the MP-7 gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a MP-7 gene; 2) an addition of one or more nucleotides to a MP-7 gene; 3) a substitution of one or more nucleotides of a MP-7 gene, 4) a chromosomal rearrangement of a MP-7 gene; 5) an alteration in the level of a messenger RNA transcript of a MP-7 gene, 6) aberrant modification of a MP-7 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a MP-7 gene, 8) a non-wild type level of a MP-7-protein, 9) allelic loss of a MP-7 gene, and 10) inappropriate post-translational modification of a MP-7-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting alterations in a MP-7 gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the MP-7-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a MP-7 gene under conditions such that hybridization and amplification of the MP-7-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated

that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878), transcriptional
5 amplification system (Kwoh, D.Y. *et al.*, 1989, Proc. Natl. Acad. Sci. USA 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.*, 1988, Bio/Technology 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection
10 schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a MP-7 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel
15 electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

20 In other embodiments, genetic mutations in MP-7 can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in MP-7 can be identified in two dimensional
25 arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the
30 characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed

of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the MP-7 gene and detect mutations by
5 comparing the sequence of the sample MP-7 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert ((1977) *PNAS* 74:560) or Sanger ((1977) *PNAS* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques*
10 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the MP-7 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA
15 or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type MP-7 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex
20 such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched
25 regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in MP-7 cDNAs obtained from samples of cells. For example, the mutY
5 enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a MP-7 sequence, e.g., a wild-type MP-7 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme,
10 and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in MP-7 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility
15 between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; and Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control MP-7 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in
20 electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the
25 basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not
30 completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is

used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective
5 primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different
10 mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of
15 interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based
20 detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence
25 of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a MP-7
30 gene.

Furthermore, any cell type or tissue in which MP-7 is expressed may be utilized in the prognostic assays described herein.

3. Monitoring of Effects During Clinical Trials

5 Monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of a MP-7 protein (e.g., modulation of cardiovascular function, an inflammatory response, e.g., T-cell activation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase MP-7 gene expression, protein levels, or upregulate
10 MP-7 activity, can be monitored in clinical trials of subjects exhibiting decreased MP-7 gene expression, protein levels, or downregulated MP-7 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease MP-7 gene expression, protein levels, or downregulate MP-7 activity, can be monitored in clinical trials of subjects exhibiting increased MP-7 gene expression, protein levels, or
15 upregulated MP-7 activity. In such clinical trials, the expression or activity of a MP-7 gene, and preferably, other genes that have been implicated in, for example, a cardiovascular disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including MP-7, that are
20 modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) which modulates MP-7 activity (e.g., identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cardiovascular disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of MP-7 and other genes implicated in a cardiovascular or immune
25 disorder. The levels of gene expression (i.e., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of MP-7 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the
30 cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a MP-7 protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the MP-7 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the MP-7 protein, mRNA, or genomic DNA in the pre-administration sample with the MP-7 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of MP-7 to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of MP-7 to lower levels than detected, i.e. to decrease the effectiveness of the agent. According to such an embodiment, MP-7 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

C. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant MP-7 expression or activity. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics.

"Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (e.g., a patient's "drug response phenotype", or "drug response genotype".) Thus, another aspect of the

invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the MP-7 molecules of the present invention or MP-7 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant MP-7 expression or activity, by administering to the subject a MP-7 or an agent which modulates MP-7 expression or at least one MP-7 activity. Subjects at risk for a disease which is caused or contributed to by aberrant MP-7 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the MP-7 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of MP-7 aberrancy, for example, a MP-7, MP-7 agonist or MP-7 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein. The prophylactic methods of the present invention are further discussed in the following subsections.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating MP-7 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a MP-7 or agent that modulates one or more of the activities of MP-7 protein activity associated with the cell. An agent that modulates MP-7 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a MP-7 protein, a MP-7 antibody, a MP-7 agonist or antagonist, a peptidomimetic of a MP-7 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more MP-7 activities. Examples of such

stimulatory agents include active MP-7 protein and a nucleic acid molecule encoding MP-7 that has been introduced into the cell. In another embodiment, the agent inhibits one or more MP-7 activities. Examples of such inhibitory agents include antisense MP-7 nucleic acid molecules, anti-MP-7 antibodies, and MP-7 inhibitors. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent), *in vivo* (e.g., by administering the agent to a subject), or alternatively *in situ* (e.g., at the site of lesion or injury, for example, in the cardiovascular system, e.g., cardiac tissue). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a MP-7 protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) MP-7 expression or activity. In another embodiment, the method involves administering a MP-7 protein or nucleic acid molecule as therapy to compensate for reduced or aberrant MP-7 expression or activity.

Stimulation of MP-7 activity is desirable in situations in which MP-7 is abnormally downregulated and/or in which increased MP-7 activity is likely to have a beneficial effect. For example, stimulation of MP-7 activity is desirable in situations in which a MP-7 is downregulated and/or in which increased MP-7 activity is likely to have a beneficial effect. Likewise, inhibition of MP-7 activity is desirable in situations in which MP-7 is abnormally upregulated and/or in which decreased MP-7 activity is likely to have a beneficial effect.

3. Pharmacogenomics

The MP-7 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on MP-7 activity (e.g., MP-7 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) immune disorders associated with aberrant MP-7 activity. In conjunction with such treatment, pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism

of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a MP-7 molecule or
5 MP-7 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a MP-7 molecule or MP-7 modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See e.g., Eichelbaum, M., *Clin Exp Pharmacol Physiol*, 1996, 23(10-11):983-
10 985 and Linder, M.W., *Clin Chem*, 1997, 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic
15 defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug
20 response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (e.g., a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically
25 significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of
30 DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-

associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among
5 such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drugs target is known (e.g., a MP-7 protein or MP-7 receptor of the present invention), all common variants of that gene can be fairly easily identified in the
10 population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT
15 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among
20 different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as
25 demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to
30 identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (e.g., a MP-7 molecule or MP-7 modulator of the present

invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a MP-7 molecule or MP-7 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

EXAMPLES

Example 1: Identification and Characterization of Human MP-7 cDNA

In this example, the identification and characterization of the gene encoding human MP-7 is described.

Isolation of the human MP-7 cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel protein, referred to herein as MP-7. The human MP-7 was isolated from a cDNA library which was prepared from tissue obtained from a subject suffering from congestive heart failure. Briefly, a cardiac tissue sample was obtained from a biopsy of a 42 year old woman suffering from congestive heart failure. mRNA was isolated from the cardiac tissue and a cDNA library was prepared therefrom using art known methods (described in, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989).

Clones were isolated, the sequence of one clone was determined and found to contain an open reading frame of 335 amino acids termed "Myocardium Protein -7" or MP-7. Signal peptide algorithms (such as the prediction program SIGNALP (Henrik Nielsen, Jacob Engelbrecht, Soren Brunak and Gunnar von Heijne "Identification of
5 prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." (1997) Protein Engineering 10, 1-6)) predict that MP-7 contains a signal peptide (amino acids 1-23 of SEQ ID NO:2). Cleavage of the putative signal peptide would result in the secretion of a 312 amino acid protein as set forth in SEQ ID NO:4 with a predicted molecular weight of approximately 34.9 kilodaltons (kD). Further, the nucleic acid
10 molecule which encodes such a MP-7 protein without the signal peptide sequence includes, for example, about 936 nucleotides of SEQ ID NO:5.

The nucleotide sequence encoding the human MP-7 protein is shown in Figure 1 and is set forth as SEQ ID NO:1. The full length protein encoded by this nucleic acid comprises about 335 amino acids and has the amino acid sequence shown in Figure 1
15 and set forth as SEQ ID NO:2. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. The clone comprising the entire coding region of human MP-7 was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on May 20, 1998, and assigned Accession No. 209887.

20

Analysis of Human MP-7

A BLAST search (Altschul *et al.* (1990) *J. Mol. Biol.* 215:403) of the nucleotide and protein sequences of human MP-7 revealed that MP-7 is similar to the following proteins: mouse antigen LY-9 precursor (Accession No. Q01965), H sapiens LY-9 gene
25 product (Accession No. L42621), H sapiens leukocyte antigen CD84 (Accession No. U82988). These proteins are approximately 27.8% identical (over MP-7 amino acids 1-326), 31.4% identical (over MP-7 amino acids 1-248), 30.2% identical (over MP-7 amino acids 1-335), to MP-7, respectively, at the amino acid level. An alignment of MP-7 and human LY-9 is shown in figure 3.

30

Tissue Distribution of MP-7 mRNA

This Example describes the tissue distribution of MP-7 mRNA, as determined by Northern blot hybridization.

Northern blot hybridizations with the various RNA samples are performed under
5 standard conditions and washed under stringent conditions, i.e., 0.2xSSC at 65°C. A
DNA probe corresponding to the coding region of MP-7 (SEQ ID NO:3) are used. The
DNA is radioactively labeled with ³²P-dCTP using the Prime-It kit (Stratagene, La
Jolla, CA) according to the instructions of the supplier. Filters containing human
mRNA (MultiTissue Northern I and MultiTissue Northern II from Clontech, Palo Alto,
10 CA) are probed in ExpressHyb hybridization solution (Clontech) and washed at high
stringency according to manufacturer's recommendations.

Example 2: Chromosomal Localization of the Human MP-7 Gene

The *MP-7* gene was mapped to human chromosome 1 by PCR typing of the
15 Genebridge (G4) radiation hybrid panel (Research Genetics, Inc., Huntsville, AL).
Typing of the DNA and comparison to radiation hybrid map data at the Whitehead
Institute Center for Genome Research (WICGR) linked the *MP-7* gene to CMD1A,
cardiomyopathy, dilated, 1A (Kass, S. *et al.*, "A gene defect that causes conduction
system disease and dilated cardiomyopathy maps chromosome 1p1-1q1," Nature Genet.
20 7:546-551. 1994, see also the CMD1A OMIM entry at
[http://www.ncbi.nlm.nih.gov/htbin-post/Omim/](http://www.ncbi.nlm.nih.gov/htbin-post/Omim/dispmim?115200)
<http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?uid=7951328&form=6&db=m&Dopt=b>).

As the panels used in the mapping studies included both human and hamster
25 sequences, the two primers to be used in the mapping of the *MP-7* gene were tested to
confirm that they were specific for human DNA rather than hamster DNA. Primers
were designed from 3' UTR sequence of *MP-7*. The *MP-7* primers used in the PCR
mapping studies were: forward GATATGACCTTCATCTGCGTTGC (SEQ ID NO:6)
and reverse CAGCAGCACCTTCACAGAGC (SEQ ID NO:7) were first tested on
30 human and hamster cell line DNA for specific amplification. Each PCR reaction
consisted of: 5µl (10ng/µl) template DNA, 1.5µl primers (6.6µM each), 1.5µl 10 X

- 95 -

Perkin Elmer PCR buffer (15mM MgCl₂, 100mM Tris-HCl, 500mM KCl Perkin-Elmer, Co., Norwalk, CT), 5u Gibco/BRL Platinum Taq polymerase (0.05u/μl Gibco/BRL Platinum Taq (Hot Start)., Norwalk, CT), and 1.2 μl Pharmacia dNTP mix (2.5 mM). Reactions were thermocycled on a Perkin-Elmer 9600 for 95°C for 10 min

5 Hot Start, 94°C 40 sec, 55°C 40 sec., 72°C, 40 sec., 35 cycles, followed by 72°C 5 minutes. Resulting PCR products were run out on a 2% agarose gel, post-stained with SYBR Gold (1:10,000 dil in 1X TBE), and scanned on a Molecular Dynamics 595 Fluorimager. The primers specifically amplified a 89 bp product from control human cell line DNA and multiple faint products from control Hamster cell line DNA. These

10 primers were used to amplify the 93 DNAs in duplicate from the Genebridge4 Radiation Hybrid Panel.

After the primers to be used in the mapping studies were determined to be specific for human DNA, the radiation hybrid mapping studies were performed as follows: PCR reactions of radiation hybrid panels, GeneBridge 4 (Research Genetics,

15 Inc., Huntsville, AL), were assembled in duplicate using an automated PCR assembly program on a Hamilton Microlab 2200 robot. Each PCR reaction consisted of: 5μl (10ng/μl) template DNA, 1.5μl primers (6.6μM each), 1.5μl 10 X Perkin Elmer PCR buffer (15mM MgCl₂, 100mM Tris-HCl, 500mM KCl Perkin- Elmer, Co., Norwalk, CT), 5u Gibco/BRL Platinum Taq polymerase (0.05u/μl Gibco/BRL Platinum Taq (Hot

20 Start)., Norwalk, CT), and 1.2 μl Pharmacia dNTP mix (2.5 mM). Reactions were thermocycled on a Perkin-Elmer 9600 95°C for 10 min Hot Start, 94°C 40 sec, 55°C 40 sec., 72°C, 40 sec., 35 cycles, followed by 72°C 5 minutes. Resulting PCR products were run out on a 2% agarose gel, post-stained with SYBR Gold (1:10,000 dil in 1X TBE), and scanned on a Molecular Dynamics 595 Fluorimager.

25 Positive hybrids for the Genebridge 4 panel were: positive:4, 6, 10,11, 13, 15, 17, 19, 26, 28, 33, 34, 38, 39, 41, 43, 44, 45, 47, 50, 52,54, 62, 63, 72, 74, 75, 79, 80, 82, 89. The following Genebridge4 hybrid DNAs were scored as questionable: 1 and 93, and the remaining DNAs were scored as negative (no human band amplified). RH linkage analysis was performed using the Map Manager QTb23 software package. MSP-7 was

30 found to map human chromosome 1, 10.8 cR₃₀₀₀ telomeric to the Whitehead Institute

framework marker AFMA323ZE5, and 19.3 cR₃₀₀₀ centromeric of the Whitehead framework marker D1S2635. LOD scores for linkage were 17.5 for AFMA323ZE5 and 14.2 for D1S2635.

5 **Example 3: Expression of Recombinant MP-7 Protein in Bacterial Cells**

In this example, MP-7 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, MP-7 is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. As the human MP-7 protein is predicted to be
10 approximately 37.4 kDa, and GST is predicted to be 26 kDa, the fusion polypeptide is predicted to be approximately 63.4 kDa, in molecular weight. Expression of the GST-MP-7 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic
15 analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

Example 4: Expression of Recombinant MP-7 Protein in COS Cells

To express the MP-7 gene in COS cells, the pcDNA/Amp vector by Invitrogen
20 Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire MP-7 protein and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the
25 polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the MP-7 DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the MP-7 coding sequence starting from the
30 initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last

20 nucleotides of the MP-7 coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the MP-7 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5a, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the MP-7-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the MP-7 polypeptide is detected by radiolabelling (^{35}S -methionine or ^{35}S -cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labelled for 8 hours with ^{35}S -methionine (or ^{35}S -cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the MP-7 coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the MP-7 polypeptide is detected by radiolabelling and immunoprecipitation using an MP-7 specific monoclonal antibody.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following

5 claims.

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:
 - (a) a nucleic acid molecule comprising the nucleotide sequence set forth
5 in SEQ ID NO:1;
 - (b) a nucleic acid molecule comprising the nucleotide sequence set forth
in SEQ ID NO:3;
 - (c) a nucleic acid molecule comprising the nucleotide sequence set forth
in SEQ ID NO:5;
- 10 2. An isolated nucleic acid molecule selected from the group consisting of a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4.
- 15 3. An isolated nucleic acid molecule selected from the group consisting of a nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number 209887.
- 20 4. An isolated nucleic acid molecule selected from the group consisting of a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4 wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3 or 5 under stringent conditions.

- 100 -

5. An isolated nucleic acid molecule selected from the group consisting of:
- a) a nucleic acid molecule comprising a nucleotide sequence which is at least 51% homologous to the nucleotide sequence of SEQ ID NO:1, 3 or 5, or a complement thereof;
 - 5 b) a nucleic acid molecule comprising a fragment of at least 107 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3 or 5, or a complement thereof;
 - c) a nucleic acid molecule which encodes a polypeptide comprising an amino acid sequence at least about 31.4% homologous to the amino acid sequence of
10 SEQ ID NO:2, SEQ ID NO:4; and
 - d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, wherein the fragment comprises at least 15 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4.
- 15
6. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 under stringent conditions.
7. An isolated nucleic acid molecule comprising a nucleotide sequence which is
20 complementary to the nucleotide sequence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.
8. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5, and a nucleotide sequence encoding a heterologous
25 polypeptide.
9. A vector comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.
- 30 10. The vector of claim 9, which is an expression vector.

11. A host cell transfected with the expression vector of claim 9.
12. A method of producing a polypeptide comprising culturing the host cell transfected with the vector of claim 9 in an appropriate culture medium to, thereby,
5 produce the polypeptide.
13. An isolated polypeptide selected from the group consisting of:
- a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, wherein the fragment comprises at least 15 contiguous
10 amino acids of SEQ ID NO:2, SEQ ID NO:4;
 - b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 3 or 5, under stringent conditions;
 - 15 c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 42.8 % homologous to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3 or 5;
 - d) a polypeptide comprising an amino acid sequence which is at least 60% homologous to the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4.
- 20
14. The isolated polypeptide of claim 13 comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4.
15. The polypeptide of claim 13, further comprising heterologous amino acid
25 sequences.
16. An antibody which selectively binds to a polypeptide of claim 13.

- 102 -

17. A method for detecting the presence of a polypeptide of claim 13 in a sample comprising:

a) contacting the sample with a compound which selectively binds to the polypeptide; and

5 b) determining whether the compound binds to the polypeptide in the sample to thereby detect the presence of a polypeptide of claim 13 in the sample.

18. The method of claim 17, wherein the compound which binds to the polypeptide is an antibody.

10

19. A kit comprising a compound which selectively binds to a polypeptide of claim 13 and instructions for use.

20. A method for detecting the presence of a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in a sample comprising:

15 a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample to thereby detect the presence of a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in the sample.

20

21. The method of claim 20, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

22. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 and instructions for use.

25

- 103 -

23. A method for identifying a compound which binds to a polypeptide of claim 13 comprising:

- a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and
- 5 b) determining whether the polypeptide binds to the test compound.

24. The method of claim 23, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

- a) detection of binding by direct detection of test
- 10 compound/polypeptide binding;
- b) detection of binding using a competition binding assay; and
- c) detection of binding using an assay for MP-7 activity.

25. A method for modulating the activity of a polypeptide of claim 13 comprising

15 contacting the polypeptide or a cell expressing the polypeptide with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

26. A method for identifying a compound which modulates the activity of a

20 polypeptide of claim 13 comprising:

- a) contacting a polypeptide of claim 13 with a test compound; and
- b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

25

1/3

10 20 30 40 50 60 70

ATGGCTGGT CCCCACATG CCTCACCTC ATCTATATCC TTTGGCAGCT CACAGGGTCA GCAGCCTCTG 70
GACCCGTGAA AGAGCTGGTC GGTTCGGTG GTGGGGCCGT GACTTTCCCC CTGAAGTCCA AAGTAAAGCA 140
AGTTBACTCT ATTGTCTGGA CCTTCAACAC AACCCTCTT GTCACCATAC AGCAGAAGG GGGCACTATC 210
ATAGTGACCC AAAATCGTAA TAGGGAGAGA GTAGACTTCC CAGATGGAGG CTACTCCCTG AAGCTCAGCA 280
AACTGAAGAA GAATGACTCA GGGATCTACT ATGTGGGGAT ATACAGCTCA TCACTCCAGC AGCCCTCCAC 350
CCAGGASTAC GTGCTGCATG TCTACGAGCA CCTGTCAAAG CCTAAAGTCA CCATGGGTCT GCAGAGCAAT 420
AAGAATGGCA CCTGTGTGAC CAATCTGACA TGCTGCATGG AACATGGGA AGAGGATGTG ATTTATACCT 480
GSAAGGCCCT GGGGCAAGCA GCCAATGAGT CCCATAATGG GTCCATCCTC CCCATCTCCT GGAGATGGGG 560
AGAAAGTGAT ATGACCTTCA TCTGCGTTGC CAGGAACCCT GTCAGCAGAA ACTTCTCAAG CCCCATCCTT 630
GCCAGGAAGC TCTGTGAAGG TGCTGCTGAT GACCCAGATT CCTCCATGGT CCTCCTGTGT CTCCTGTTGG 700
TGCCCTCCT GCTCAGTCTC TTTGTACTGG GGCTATTTCT TTGTTTCTG AAGAGAGAGA GACAAGAAGA 770
GTACATTGAA GAGAAGAAGA GAGTGGACAT TTGTCGGGAA ACTCCTAACA TATGCCCCCA TTCTGGAGAG 840
AACACAGAGT ACGACACAAT CCCTCACACT AATAGAACAA TCCTAAAGGA AGATCCAGCA AATACGGTTT 910
ACTCCACTGT GGAATACCG AAAAGATGG AAAATCCCCA CTCATGCTC ACGATGCCAG ACACACCAAG 980
GCTATTTGCC TATGAGAATG TTATCTAG 1008

FIGURE 1

2/3

10 20 30 40 50 60 70
:IAGSPTCLTLIYILWOLTGSAASGPVKELVGSVGGAYTFPLKSKVKQVDSI VTFNTTPLYTIQPEGGTI 70
IVTONRRNRERVD FPDGGYSLKLSKLKKNDSGIYYVG IYSSSLQDPSTQ EYVLHYEHLSPKVTHGLQSN 140
KNGTCVTNL TCCNEHGEEDVIYTKALGQAANESHNGS ILPI SWRWGESDHTFICVARNPVSRNFSSPIL 210
ARKLCEGAADDPDSSHVLLCLLLVPLLLSLFVLGLFLWFLKREROEYIEEKKRVDICRETPNICPHSGE 280
NTEYDTIPHTNRTILKEDPANTVYSTVEIPKKHENPHSLLTHPOTPRLFAYENVI. 336

FIGURE 2

3/3

Seq1(1>335)	Seq2(1>610)	Similarity	Gap	Gap	Consensus	
MSP-7	LY-2 (121569)	Index	Number	Length	Length	
(18>248)	(216>442)	31.4	5	11	234	
↖20	↖30	↖40	↖50	↖60	↖70	↖80
GSAASGPVKELVGSVGGAVTFPLK-SKVKQVDSIVWTFNTTPLVTIQPEGGT---IIVTONRNRERVDFF						
:S :... :VG :G.:VT:PL : :...:VW FNT: : : E::T :I :... :RV :						
ASRGGTTGETVVGVLGEPVTLPLALPACRDTEKVVWLFNTSIIKEREAAATADPLIKSRDPYKNRVWVS						
↖220	↖230	↖240	↖250	↖260	↖270	↖280
↖90	↖100	↖110	↖120	↖130	↖140	↖150
DGGYSLKLSKLKKNDSGIYYVGIYSSSLQOPSTOEYVLHVYEHLSKPKVTMGLOSNGTCVTNLTCME						
: :SLK:S:LK :D:G Y...:S.: :S :.L :Y :L.KPK:T :L: :.:G.C :LTC.:E						
SQDCSLKISQLKIEDAGPYHAYVCSEASSVTSMTHTVLLIYRRLRKPKITWSLRHSEDGICRISLTCSVE						
↖290	↖300	↖310	↖320	↖330	↖340	↖350
↖160	↖170	↖180	↖190	↖200	↖210	↖220
HGEEDVIYTWKALGQAANESHNGSILPISWRWGESDHTFICVARNPVSRNFSSPILARKLCEGAADDPDS						
:G.:V:YTW.:L :.A S...S L :SWR :E:: :.C.A.NPVSR: S :.L: :.:C.G P..						
DGGNTVNYTWTPLQKEAVVSQGESHLNVSWRSSENHPNLTCTASNPVSR-SHQFLSENICSG----PER						
↖360	↖370	↖380	↖390	↖400	↖410	↖420
↖230	↖240					
SHVLLCLLLVPLLLSLFVLGLFLW						
: L L:: :.L: :G:F W						
NTKLWIGLFL--MVCLLCVGIFSW						
↖430	↖440					

FIGURE 3

- 1 -

SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc.

<120> NOVEL MP-7 PROTEIN AND NUCLEIC ACID MOLECULES AND USES
THEREFOR

<130> MNI-048CPPC

<140>

<141>

<150> USSN 09/261,759

<151> 1999-03-02

<150> USSN 09/163,284

<151> 1998-09-29

<150> USSN 60/090,579

<151> 1998-06-25

<160> 7

<170> PatentIn Ver. 2.0

<210> 1

<211> 1008

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1005)

<400> 1

atg gct ggt tcc cca aca tgc ctc acc ctc atc tat atc ctt tgg cag	48
Met Ala Gly Ser Pro Thr Cys Leu Thr Leu Ile Tyr Ile Leu Trp Gln	
1 5 10 15	
ctc aca ggg tca gca gcc tct gga ccc gtg aaa gag ctg gtc ggt tcc	96
Leu Thr Gly Ser Ala Ala Ser Gly Pro Val Lys Glu Leu Val Gly Ser	
20 25 30	
gtt ggt ggg gcc gtg act ttc ccc ctg aag tcc aaa gta aag caa gtt	144
Val Gly Gly Ala Val Thr Phe Pro Leu Lys Ser Lys Val Lys Gln Val	
35 40 45	
gac tct att gtc tgg acc ttc aac aca acc cct ctt gtc acc ata cag	192
Asp Ser Ile Val Trp Thr Phe Asn Thr Thr Pro Leu Val Thr Ile Gln	
50 55 60	
cca gaa ggg ggc act atc ata gtg acc caa aat cgt aat agg gag aga	240
Pro Glu Gly Gly Thr Ile Ile Val Thr Gln Asn Arg Asn Arg Glu Arg	
65 70 75 80	
gta gac ttc cca gat gga ggc tac tcc ctg aag ctc agc aaa ctg aag	288
Val Asp Phe Pro Asp Gly Gly Tyr Ser Leu Lys Leu Ser Lys Leu Lys	
85 90 95	

- 2 -

aag aat gac tca ggg atc tac tat gtg ggg ata tac agc tca tca ctc	336
Lys Asn Asp Ser Gly Ile Tyr Tyr Val Gly Ile Tyr Ser Ser Ser Leu	
100 105 110	
cag cag ccc tcc acc cag gag tac gtg ctg cat gtc tac gag cac ctg	384
Gln Gln Pro Ser Thr Gln Glu Tyr Val Leu His Val Tyr Glu His Leu	
115 120 125	
tca aag cct aaa gtc acc atg ggt ctg cag agc aat aag aat ggc acc	432
Ser Lys Pro Lys Val Thr Met Gly Leu Gln Ser Asn Lys Asn Gly Thr	
130 135 140	
tgt gtg acc aat ctg aca tgc tgc atg gaa cat ggg gaa gag gat gtg	480
Cys Val Thr Asn Leu Thr Cys Cys Met Glu His Gly Glu Glu Asp Val	
145 150 155 160	
att tat acc tgg aag gcc ctg ggg caa gca gcc aat gag tcc cat aat	528
Ile Tyr Thr Trp Lys Ala Leu Gly Gln Ala Ala Asn Glu Ser His Asn	
165 170 175	
ggg tcc atc ctc ccc atc tcc tgg aga tgg gga gaa agt gat atg acc	576
Gly Ser Ile Leu Pro Ile Ser Trp Arg Trp Gly Glu Ser Asp Met Thr	
180 185 190	
ttc atc tgc gtt gcc agg aac cct gtc agc aga aac ttc tca agc ccc	624
Phe Ile Cys Val Ala Arg Asn Pro Val Ser Arg Asn Phe Ser Ser Pro	
195 200 205	
atc ctt gcc agg aag ctc tgt gaa ggt gct gct gat gac cca gat tcc	672
Ile Leu Ala Arg Lys Leu Cys Glu Gly Ala Ala Asp Asp Pro Asp Ser	
210 215 220	
tcc atg gtc ctc ctg tgt ctc ctg ttg gtg ccc ctc ctg ctc agt ctc	720
Ser Met Val Leu Leu Cys Leu Leu Leu Val Pro Leu Leu Leu Ser Leu	
225 230 235 240	
ttt gta ctg ggg cta ttt ctt tgg ttt ctg aag aga gag aga caa gaa	768
Phe Val Leu Gly Leu Phe Leu Trp Phe Leu Lys Arg Glu Arg Gln Glu	
245 250 255	
gag tac att gaa gag aag aag aga gtg gac att tgt cgg gaa act cct	816
Glu Tyr Ile Glu Glu Lys Lys Arg Val Asp Ile Cys Arg Glu Thr Pro	
260 265 270	
aac ata tgc ccc cat tct gga gag aac aca gag tac gac aca atc cct	864
Asn Ile Cys Pro His Ser Gly Glu Asn Thr Glu Tyr Asp Thr Ile Pro	
275 280 285	
cac act aat aga aca atc cta aag gaa gat cca gca aat acg gtt tac	912
His Thr Asn Arg Thr Ile Leu Lys Glu Asp Pro Ala Asn Thr Val Tyr	
290 295 300	
tcc act gtg gaa ata ccg aaa aag atg gaa aat ccc cac tca ctg ctc	960
Ser Thr Val Glu Ile Pro Lys Lys Met Glu Asn Pro His Ser Leu Leu	
305 310 315 320	
acg atg cca gac aca cca agg cta ttt gcc tat gag aat gtt atc tag	1008
Thr Met Pro Asp Thr Pro Arg Leu Phe Ala Tyr Glu Asn Val Ile	
325 330 335	

- 3 -

<210> 2

<211> 335

<212> PRT

<213> Homo sapiens

<400> 2

Met Ala Gly Ser Pro Thr Cys Leu Thr Leu Ile Tyr Ile Leu Trp Gln
 1 5 10 15

Leu Thr Gly Ser Ala Ala Ser Gly Pro Val Lys Glu Leu Val Gly Ser
 20 25 30

Val Gly Gly Ala Val Thr Phe Pro Leu Lys Ser Lys Val Lys Gln Val
 35 40 45

Asp Ser Ile Val Trp Thr Phe Asn Thr Thr Pro Leu Val Thr Ile Gln
 50 55 60

Pro Glu Gly Gly Thr Ile Ile Val Thr Gln Asn Arg Asn Arg Glu Arg
 65 70 75 80

Val Asp Phe Pro Asp Gly Gly Tyr Ser Leu Lys Leu Ser Lys Leu Lys
 85 90 95

Lys Asn Asp Ser Gly Ile Tyr Tyr Val Gly Ile Tyr Ser Ser Ser Leu
 100 105 110

Gln Gln Pro Ser Thr Gln Glu Tyr Val Leu His Val Tyr Glu His Leu
 115 120 125

Ser Lys Pro Lys Val Thr Met Gly Leu Gln Ser Asn Lys Asn Gly Thr
 130 135 140

Cys Val Thr Asn Leu Thr Cys Cys Met Glu His Gly Glu Glu Asp Val
 145 150 155 160

Ile Tyr Thr Trp Lys Ala Leu Gly Gln Ala Ala Asn Glu Ser His Asn
 165 170 175

Gly Ser Ile Leu Pro Ile Ser Trp Arg Trp Gly Glu Ser Asp Met Thr
 180 185 190

Phe Ile Cys Val Ala Arg Asn Pro Val Ser Arg Asn Phe Ser Ser Pro
 195 200 205

Ile Leu Ala Arg Lys Leu Cys Glu Gly Ala Ala Asp Asp Pro Asp Ser
 210 215 220

Ser Met Val Leu Leu Cys Leu Leu Leu Val Pro Leu Leu Leu Ser Leu
 225 230 235 240

Phe Val Leu Gly Leu Phe Leu Trp Phe Leu Lys Arg Glu Arg Gln Glu
 245 250 255

Glu Tyr Ile Glu Glu Lys Lys Arg Val Asp Ile Cys Arg Glu Thr Pro
 260 265 270

- 4 -

Asn Ile Cys Pro His Ser Gly Glu Asn Thr Glu Tyr Asp Thr Ile Pro
 275 280 285

His Thr Asn Arg Thr Ile Leu Lys Glu Asp Pro Ala Asn Thr Val Tyr
 290 295 300

Ser Thr Val Glu Ile Pro Lys Lys Met Glu Asn Pro His Ser Leu Leu
 305 310 315 320

Thr Met Pro Asp Thr Pro Arg Leu Phe Ala Tyr Glu Asn Val Ile
 325 330 335

<210> 3
 <211> 1005
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1005)

<400> 3
 atg gct ggt tcc cca aca tgc ctc acc ctc atc tat atc ctt tgg cag 48
 Met Ala Gly Ser Pro Thr Cys Leu Thr Leu Ile Tyr Ile Leu Trp Gln
 1 5 10 15

ctc aca ggg tca gca gcc tct gga ccc gtg aaa gag ctg gtc ggt tcc 96
 Leu Thr Gly Ser Ala Ala Ser Gly Pro Val Lys Glu Leu Val Gly Ser
 20 25 30

gtt ggt ggg gcc gtg act ttc ccc ctg aag tcc aaa gta aag caa gtt 144
 Val Gly Gly Ala Val Thr Phe Pro Leu Lys Ser Lys Val Lys Gln Val
 35 40 45

gac tct att gtc tgg acc ttc aac aca acc cct ctt gtc acc ata cag 192
 Asp Ser Ile Val Trp Thr Phe Asn Thr Thr Pro Leu Val Thr Ile Gln
 50 55 60

cca gaa ggg ggc act atc ata gtg acc caa aat cgt aat agg gag aga 240
 Pro Glu Gly Gly Thr Ile Ile Val Thr Gln Asn Arg Asn Arg Glu Arg
 65 70 75 80

gta gac ttc cca gat gga ggc tac tcc ctg aag ctc agc aaa ctg aag 288
 Val Asp Phe Pro Asp Gly Gly Tyr Ser Leu Lys Leu Ser Lys Leu Lys
 85 90 95

aag aat gac tca ggg atc tac tat gtg ggg ata tac agc tca tca ctc 336
 Lys Asn Asp Ser Gly Ile Tyr Tyr Val Gly Ile Tyr Ser Ser Ser Leu
 100 105 110

cag cag ccc tcc acc cag gag tac gtg ctg cat gtc tac gag cac ctg 384
 Gln Gln Pro Ser Thr Gln Glu Tyr Val Leu His Val Tyr Glu His Leu
 115 120 125

tca aag cct aaa gtc acc atg ggt ctg cag agc aat aag aat ggc acc 432
 Ser Lys Pro Lys Val Thr Met Gly Leu Gln Ser Asn Lys Asn Gly Thr
 130 135 140

- 5 -

tgt gtg acc aat ctg aca tgc tgc atg gaa cat ggg gaa gag gat gtg 480
 Cys Val Thr Asn Leu Thr Cys Cys Met Glu His Gly Glu Glu Asp Val
 145 150 155 160

att tat acc tgg aag gcc ctg ggg caa gca gcc aat gag tcc cat aat 528
 Ile Tyr Thr Trp Lys Ala Leu Gly Gln Ala Ala Asn Glu Ser His Asn
 165 170 175

ggg tcc atc ctc ccc atc tcc tgg aga tgg gga gaa agt gat atg acc 576
 Gly Ser Ile Leu Pro Ile Ser Trp Arg Trp Gly Glu Ser Asp Met Thr
 180 185 190

ttc atc tgc gtt gcc agg aac cct gtc agc aga aac ttc tca agc ccc 624
 Phe Ile Cys Val Ala Arg Asn Pro Val Ser Arg Asn Phe Ser Ser Pro
 195 200 205

atc ctt gcc agg aag ctc tgt gaa ggt gct gct gat gac cca gat tcc 672
 Ile Leu Ala Arg Lys Leu Cys Glu Gly Ala Ala Asp Asp Pro Asp Ser
 210 215 220

tcc atg gtc ctc ctg tgt ctc ctg ttg gtg ccc ctc ctg ctc agt ctc 720
 Ser Met Val Leu Leu Cys Leu Leu Leu Val Pro Leu Leu Leu Ser Leu
 225 230 235 240

ttt gta ctg ggg cta ttt ctt tgg ttt ctg aag aga gag aga caa gaa 768
 Phe Val Leu Gly Leu Phe Leu Trp Phe Leu Lys Arg Glu Arg Gln Glu
 245 250 255

gag tac att gaa gag aag aag aga gtg gac att tgt cgg gaa act cct 816
 Glu Tyr Ile Glu Glu Lys Lys Arg Val Asp Ile Cys Arg Glu Thr Pro
 260 265 270

aac ata tgc ccc cat tct gga gag aac aca gag tac gac aca atc cct 864
 Asn Ile Cys Pro His Ser Gly Glu Asn Thr Glu Tyr Asp Thr Ile Pro
 275 280 285

cac act aat aga aca atc cta aag gaa gat cca gca aat acg gtt tac 912
 His Thr Asn Arg Thr Ile Leu Lys Glu Asp Pro Ala Asn Thr Val Tyr
 290 295 300

tcc act gtg gaa ata ccg aaa aag atg gaa aat ccc cac tca ctg ctc 960
 Ser Thr Val Glu Ile Pro Lys Lys Met Glu Asn Pro His Ser Leu Leu
 305 310 315 320

acg atg cca gac aca cca agg cta ttt gcc tat gag aat gtt atc 1005
 Thr Met Pro Asp Thr Pro Arg Leu Phe Ala Tyr Glu Asn Val Ile
 325 330 335

<210> 4

<211> 335

<212> PRT

<213> Homo sapiens

<400> 4

Met Ala Gly Ser Pro Thr Cys Leu Thr Leu Ile Tyr Ile Leu Trp Gln
 1 5 10 15

Leu Thr Gly Ser Ala Ala Ser Gly Pro Val Lys Glu Leu Val Gly Ser

- 6 -

20					25					30					
Val	Gly	Gly	Ala	Val	Thr	Phe	Pro	Leu	Lys	Ser	Lys	Val	Lys	Gln	Val
		35					40					45			
Asp	Ser	Ile	Val	Trp	Thr	Phe	Asn	Thr	Thr	Pro	Leu	Val	Thr	Ile	Gln
	50					55					60				
Pro	Glu	Gly	Gly	Thr	Ile	Ile	Val	Thr	Gln	Asn	Arg	Asn	Arg	Glu	Arg
	65					70					75				80
Val	Asp	Phe	Pro	Asp	Gly	Gly	Tyr	Ser	Leu	Lys	Leu	Ser	Lys	Leu	Lys
				85					90					95	
Lys	Asn	Asp	Ser	Gly	Ile	Tyr	Tyr	Val	Gly	Ile	Tyr	Ser	Ser	Ser	Leu
			100					105					110		
Gln	Gln	Pro	Ser	Thr	Gln	Glu	Tyr	Val	Leu	His	Val	Tyr	Glu	His	Leu
		115					120					125			
Ser	Lys	Pro	Lys	Val	Thr	Met	Gly	Leu	Gln	Ser	Asn	Lys	Asn	Gly	Thr
	130					135					140				
Cys	Val	Thr	Asn	Leu	Thr	Cys	Cys	Met	Glu	His	Gly	Glu	Glu	Asp	Val
	145					150					155				160
Ile	Tyr	Thr	Trp	Lys	Ala	Leu	Gly	Gln	Ala	Ala	Asn	Glu	Ser	His	Asn
				165					170					175	
Gly	Ser	Ile	Leu	Pro	Ile	Ser	Trp	Arg	Trp	Gly	Glu	Ser	Asp	Met	Thr
			180					185					190		
Phe	Ile	Cys	Val	Ala	Arg	Asn	Pro	Val	Ser	Arg	Asn	Phe	Ser	Ser	Pro
		195					200					205			
Ile	Leu	Ala	Arg	Lys	Leu	Cys	Glu	Gly	Ala	Ala	Asp	Asp	Pro	Asp	Ser
	210					215					220				
Ser	Met	Val	Leu	Leu	Cys	Leu	Leu	Leu	Val	Pro	Leu	Leu	Leu	Ser	Leu
	225					230					235				240
Phe	Val	Leu	Gly	Leu	Phe	Leu	Trp	Phe	Leu	Lys	Arg	Glu	Arg	Gln	Glu
				245					250					255	
Glu	Tyr	Ile	Glu	Glu	Lys	Lys	Arg	Val	Asp	Ile	Cys	Arg	Glu	Thr	Pro
			260					265					270		
Asn	Ile	Cys	Pro	His	Ser	Gly	Glu	Asn	Thr	Glu	Tyr	Asp	Thr	Ile	Pro
		275					280					285			
His	Thr	Asn	Arg	Thr	Ile	Leu	Lys	Glu	Asp	Pro	Ala	Asn	Thr	Val	Tyr
		290					295					300			
Ser	Thr	Val	Glu	Ile	Pro	Lys	Lys	Met	Glu	Asn	Pro	His	Ser	Leu	Leu
	305					310					315				320
Thr	Met	Pro	Asp	Thr	Pro	Arg	Leu	Phe	Ala	Tyr	Glu	Asn	Val	Ile	
				325					330					335	

- 7 -

<210> 5
 <211> 936
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(936)

<400> 5
 gga ccc gtg aaa gag ctg gtc ggt tcc gtt ggt ggg gcc gtg act ttc 48
 Gly Pro Val Lys Glu Leu Val Gly Ser Val Gly Gly Ala Val Thr Phe
 1 5 10 15

ccc ctg aag tcc aaa gta aag caa gtt gac tct att gtc tgg acc ttc 96
 Pro Leu Lys Ser Lys Val Lys Gln Val Asp Ser Ile Val Trp Thr Phe
 20 25 30

aac aca acc cct ctt gtc acc ata cag cca gaa ggg ggc act atc ata 144
 Asn Thr Thr Pro Leu Val Thr Ile Gln Pro Glu Gly Gly Thr Ile Ile
 35 40 45

gtg acc caa aat cgt aat agg gag aga gta gac ttc cca gat gga ggc 192
 Val Thr Gln Asn Arg Asn Arg Glu Arg Val Asp Phe Pro Asp Gly Gly
 50 55 60

tac tcc ctg aag ctc agc aaa ctg aag aag aat gac tca ggg atc tac 240
 Tyr Ser Leu Lys Leu Ser Lys Leu Lys Lys Asn Asp Ser Gly Ile Tyr
 65 70 75 80

tat gtg ggg ata tac agc tca tca ctc cag cag ccc tcc acc cag gag 288
 Tyr Val Gly Ile Tyr Ser Ser Ser Leu Gln Gln Pro Ser Thr Gln Glu
 85 90 95

tac gtg ctg cat gtc tac gag cac ctg tca aag cct aaa gtc acc atg 336
 Tyr Val Leu His Val Tyr Glu His Leu Ser Lys Pro Lys Val Thr Met
 100 105 110

ggt ctg cag agc aat aag aat ggc acc tgt gtg acc aat ctg aca tgc 384
 Gly Leu Gln Ser Asn Lys Asn Gly Thr Cys Val Thr Asn Leu Thr Cys
 115 120 125

tgc atg gaa cat ggg gaa gag gat gtg att tat acc tgg aag gcc ctg 432
 Cys Met Glu His Gly Glu Glu Asp Val Ile Tyr Thr Trp Lys Ala Leu
 130 135 140

ggg caa gca gcc aat gag tcc cat aat ggg tcc atc ctc ccc atc tcc 480
 Gly Gln Ala Ala Asn Glu Ser His Asn Gly Ser Ile Leu Pro Ile Ser
 145 150 155 160

tgg aga tgg gga gaa agt gat atg acc ttc atc tgc gtt gcc agg aac 528
 Trp Arg Trp Gly Glu Ser Asp Met Thr Phe Ile Cys Val Ala Arg Asn
 165 170 175

cct gtc agc aga aac ttc tca agc ccc atc ctt gcc agg aag ctc tgt 576
 Pro Val Ser Arg Asn Phe Ser Ser Pro Ile Leu Ala Arg Lys Leu Cys
 180 185 190

- 8 -

gaa ggt gct gct gat gac cca gat tcc tcc atg gtc ctc ctg tgt ctc	624
Glu Gly Ala Ala Asp Asp Pro Asp Ser Ser Met Val Leu Leu Cys Leu	
195 200 205	
ctg ttg gtg ccc ctc ctg ctc agt ctc ttt gta ctg ggg cta ttt ctt	672
Leu Leu Val Pro Leu Leu Leu Ser Leu Phe Val Leu Gly Leu Phe Leu	
210 215 220	
tgg ttt ctg aag aga gag aga caa gaa gag tac att gaa gag aag aag	720
Trp Phe Leu Lys Arg Glu Arg Gln Glu Glu Tyr Ile Glu Glu Lys Lys	
225 230 235 240	
aga gtg gac att tgt cgg gaa act cct aac ata tgc ccc cat tct gga	768
Arg Val Asp Ile Cys Arg Glu Thr Pro Asn Ile Cys Pro His Ser Gly	
245 250 255	
gag aac aca gag tac gac aca atc cct cac act aat aga aca atc cta	816
Glu Asn Thr Glu Tyr Asp Thr Ile Pro His Thr Asn Arg Thr Ile Leu	
260 265 270	
aag gaa gat cca gca aat acg gtt tac tcc act gtg gaa ata ccg aaa	864
Lys Glu Asp Pro Ala Asn Thr Val Tyr Ser Thr Val Glu Ile Pro Lys	
275 280 285	
aag atg gaa aat ccc cac tca ctg ctc acg atg cca gac aca cca agg	912
Lys Met Glu Asn Pro His Ser Leu Leu Thr Met Pro Asp Thr Pro Arg	
290 295 300	
cta ttt gcc tat gag aat gtt atc	936
Leu Phe Ala Tyr Glu Asn Val Ile	
305 310	

<210> 6

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 6

gatatgacct tcattctgcgt tgc

23

<210> 7

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 7

cagcagcacc ttcacagagc

20